GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:04:17 ; Search time 24 Seconds
(without alignments)
296.681 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1252	99.1	415	4	US-09-545-944-2	Sequence 2, Appli
2	304	24.1	353	1	US-08-118-270-45	Sequence 45, Appl
3	304	24.1	353	5	PCT-US93-08528-45	Sequence 45, Appl
4	281	22.2	416	3	US-08-858-876A-4	Sequence 4, Appli
5	281	22.2	416	4	US-09-472-880-4	Sequence 4, Appli
6	277	21.9	289	4	US-09-077-675A-10	Sequence 10, Appl
7	277	21.9	361	4	US-09-077-675A-8	Sequence 8, Appli
8	277	21.9	366	4	US-09-077-675A-13	Sequence 13, Appl
9	274	21.7	353	4	US-09-077-675A-3	Sequence 3, Appli
10	274	21.7	364	4	US-09-077-675A-16	Sequence 16, Appl
11	273	21.6	289	4	US-09-077-675A-5	Sequence 5, Appli
12	265	21.0	393	1	US-07-629-104I-3	Sequence 3, Appli
13	262.5	20.8	259	4	US-09-261-599B-3	Sequence 3, Appli
14	262.5	20.8	259	4	US-09-456-455A-3	Sequence 3, Appli
15	259	20.5	398	2	US-08-288-663A-1	Sequence 1, Appli
16	258	20.4	410	3	US-08-858-876A-2	Sequence 2, Appli
17	258	20.4	410	4	US-09-472-880-2	Sequence 2, Appli
18	256.5	20.3	400	4	US-09-351-198-2	Sequence 2, Appli
19	256.5	20.3	400	4	US-09-113-426-2	Sequence 2, Appli
20	256.5	20.3	415	4	US-08-405-271A-20	Sequence 20, Appl
21	255.5	20.2	400	4	US-08-188-275A-2	Sequence 2, Appli
22	253.5	20.1	400	3	US-08-889-108-8	Sequence 8, Appli
23	253.5	20.1	400	5	PCT-US94-10358-8	Sequence 8, Appli
24	252.5	20.0	369	1	US-07-816-283-6	Sequence 6, Appli
25	252.5	20.0	369	1	US-08-417-103-6	Sequence 6, Appli
26	252.5	20.0	369	1	US-08-417-103-16	Sequence 16, Appl
27	247.5	19.6	369	2	US-08-411-859-3	Sequence 3, Appli
28	247.5	19.6	369	4	US-08-387-707-9	Sequence 9, Appli
29	247.5	19.6	369	4	US-08-405-271A-9	Sequence 9, Appli
30	247.5	19.6	398	3	US-09-170-331-5	Sequence 5, Appli
31	246	19.5	410	4	US-09-200-090-2	Sequence 2, Appli
32	245.5	19.4	369	4	US-08-120-601B-9	Sequence 9, Appli
33	242.5	19.2	356	4	US-08-430-286A-2	Sequence 2, Appli
34	242.5	19.2	356	4	US-08-430-286A-5	Sequence 5, Appli
35	242.5	19.2	367	4	US-08-405-271A-23	Sequence 23, Appl
36	242.5	19.2	391	2	US-08-454-549-3	Sequence 3, Appli
37	242.5	19.2	391	3	US-08-454-552-3	Sequence 3, Appli
38	242.5	19.2	398	1	US-08-149-093A-5	Sequence 5, Appli
39	242.5	19.2	398	2	US-08-911-245-5	Sequence 5, Appli
40	242.5	19.2	398	3	US-08-889-108-2	Sequence 2, Appli
41	242.5	19.2	398	4	US-08-120-601B-2	Sequence 2, Appli
42	242.5	19.2	398	4	US-08-188-275A-3	Sequence 3, Appli
43	242.5	19.2	398	4	US-08-387-707-16	Sequence 16, Appl
44	242.5	19.2	398	4	US-09-510-473-5	Sequence 5, Appli
45	242.5	19.2	398	4	US-09-351-198-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-545-944-2
; Sequence 2, Application US/09545944
; Patent No. 6461836
; GENERAL INFORMATION:
; APPLICANT: AMES, ROBERT
; APPLICANT: ELSHOURBAGY, NABIL
; APPLICANT: MICHALOVICH, DAVID
; APPLICANT: SARAU, HENRY
; APPLICANT: SHABON, USMAN
; APPLICANT: VAWTER, LISA
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
; TITLE OF INVENTION: (AXOR34) AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP70657-1
; CURRENT APPLICATION NUMBER: US/09/545,944
; CURRENT FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 09/435,384
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-545-944-2

Query Match 99.1%; Score 1252; DB 4; Length 415;
Best Local Similarity 99.6%; Pred. No. 3e-107;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243
```

RESULT 2
US-08-118-270-45
; Sequence 45, Application US/08118270

```
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  353 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-45

  Query Match             24.1%;  Score 304;  DB 1;  Length 353;
  Best Local Similarity   33.3%;  Pred. No. 1.9e-20;
  Matches   68;  Conservative   49;  Mismatches   73;  Indels   14;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY 101
           |: :|: :|||| :|| :    : :    |:::  :|:| |||:||||:||    :|:|
Db       4 VTAIYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHLSSLALSDLLILLW---VELY 60

Qy     102 EMWRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRR 158
             ::|: ||  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :|
Db      61 NFIWHHPWAFGDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSRS 117

Qy     159 RALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS 218
           |  : :  :|  | | ::|     |::          ||   || |        :||| :
Db     118 RTKKFISAIWLASALLAIPMLFTLGLQNR--SGDGTHPGGLVCTPIVDTATVKVVIQVNT 175

Qy     219 FLFYLLPMTVISVLYYLMALRVSI 242
           |: :| || |||:|  ::| ::::
Db     176 FMSFLFPMLVISILNTVIANKLTV 199


RESULT 3
PCT-US93-08528-45
; Sequence 45, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  353 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-45

  Query Match             24.1%;  Score 304;  DB 5;  Length 353;
  Best Local Similarity   33.3%;  Pred. No. 1.9e-20;
  Matches   68;  Conservative   49;  Mismatches   73;  Indels   14;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY 101
           |: :|: :|||| :|| :    : :    |:::  :|:| |||:||||:||    :|:|
Db       4 VTAIYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHLSSLALSDLLILLW---VELY 60

Qy     102 EMWRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRR 158
             ::|: ||  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :|
Db      61 NFIWHHPWAFGDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSRS 117

Qy     159 RALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS 218
           |  : :  :|  | | ::|     |::          ||   || |        :||| :
Db     118 RTKKFISAIWLASALLAIPMLFTLGLQNR--SGDGTHPGGLVCTPIVDTATVKVVIQVNT 175

Qy     219 FLFYLLPMTVISVLYYLMALRVSI 242
           |: :| || |||:|  ::| ::::
Db     176 FMSFLFPMLVISILNTVIANKLTV 199


RESULT 4
US-08-858-876A-4
; Sequence 4, Application US/08858876A
; Patent No. 6022856
;  GENERAL INFORMATION:
;    APPLICANT:  Daniel CAPUT
;    APPLICANT:  Pascale CHALON
;    APPLICANT:  Pascual FERRARA
;    APPLICANT:  Vita NATALIO
;    TITLE OF INVENTION:  Type 2 Neurotensin Receptor
;    TITLE OF INVENTION:  (hNT-R2)
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jacobson, Price, Holman & Stern, PLLC
;      STREET:  400 Seventh Street
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/858,876A
;      FILING DATE:  19-SEP-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR 9723204
;      FILING DATE:  17-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Player, William E.
;      REGISTRATION NUMBER:  31,049
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  416 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-858-876A-4

  Query Match             22.2%;  Score 281;  DB 3;  Length 416;
  Best Local Similarity   35.3%;  Pred. No. 3e-18;
  Matches   72;  Conservative   40;  Mismatches   78;  Indels   14;  Gaps    6;

Qy      46 SVVYVPIFVVGVIGNVLVCLVILQHQAMKT-PTNYYLFSLAVSDLLVLLLGMPLEVYE-M 103
           : :|  ||  |  || |   |:|: :| :     |:: |||:| ||:||: ||:|:|  :
Db      36 TALYSLIFAFGTAGNALSVHVVLKARAGRPGRLRYHVLSLALSALLLLVSMPMELYNFV 95

Qy     104 WRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA 160
           | :||::|| :||   ||   : |   :|::||: ::| || :|:  | ||:   | ||
Db      96 WSHYPWVFGDLGCRGYYF---VRELCAYATVLSVASLSAERCLAVCQPLRARRLLTPRRT 152

Qy     161 LRILGIVWGFSVLFSLPNTSIHGIKFHY-FPNGSLVPGSATCTVIKPMWIYNFIIQVTSF 219
            |:| :||  |:  :||   | | |      :|   | |  |||:          |||
Db     153 RRLLSLVWVASLGLALPMAVIMGQKHEVESADGEPEPASRVCTVLVSRATLQVFIQVNVL 212

Qy     220 LFYLLPMTVISVL-----YYLMAL 238
           : : ||: : : |      :||||
Db     213 VSFALPLALTAFLNGITVNHLMAL 236


RESULT 5
US-09-472-880-4
 Sequence 4, Application US/09472880
 Patent No. 6274333
  GENERAL INFORMATION:
;       APPLICANT: Daniel CAPUT
;                  Pascale CHALON
;                  Pascual FERRARA
;                  Vita NATALIO
;       TITLE OF INVENTION: Type 2 Neurotensin Receptor
;                           (hNT-R2)
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
;            STREET: 400 Seventh Street
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/472,880
;            FILING DATE: 28-Dec-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/FR 9723204
;            FILING DATE: 17-MAR-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Player, William E.
;            REGISTRATION NUMBER: 31,049
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 416 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-472-880-4

  Query Match             22.2%;  Score 281;  DB 4;  Length 416;
  Best Local Similarity   35.3%;  Pred. No. 3e-18;
  Matches   72;  Conservative   40;  Mismatches   78;  Indels   14;  Gaps    6;

Qy      46 SVVYVPIFVVGVIGNVLVCLVILQHQAMKT-PTNYYLFSLAVSDLLVLLLGMPLEVYE-M 103
           : :|  ||  |  || |   |:|: :| :     |:: |||:| ||:||: ||:|:|  :
Db      36 TALYSLIFAFGTAGNALSVHVVLKARAGRPGRLRYHVLSLALSALLLLVSMPMELYNFV 95

Qy     104 WRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA 160
           | :||::|| :||   ||   : |   :|::||: ::| || :|:  | ||:   | ||
Db      96 WSHYPWVFGDLGCRGYYF---VRELCAYATVLSVASLSAERCLAVCQPLRARRLLTPRRT 152

Qy     161 LRILGIVWGFSVLFSLPNTSIHGIKFHY-FPNGSLVPGSATCTVIKPMWIYNFIIQVTSF 219
            |:| :||  |:  :||   | | |      :|   | |  |||:          |||
Db     153 RRLLSLVWVASLGLALPMAVIMGQKHEVESADGEPEPASRVCTVLVSRATLQVFIQVNVL 212

Qy     220 LFYLLPMTVISVL-----YYLMAL 238
           : : ||: : : |      :||||
Db     213 VSFALPLALTAFLNGITVNHLMAL 236


RESULT 6
US-09-077-675A-10
; Sequence 10, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
```

```
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-077-675A-10

  Query Match             21.9%;  Score 277;  DB 4;  Length 289;
  Best Local Similarity   31.7%;  Pred. No. 4.6e-18;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : ::| || || |:| |||| |  |||::  :|
Db      46 VTATCVALFVVGIAGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 104

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     105 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 164

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
            ::|  :   : |   : |::     ||:               |   ||:  :|:
Db     165 FVIWAVAFCSAGPIFVLVGVEHE---NGTDPWDTNECRPTEFAVRSGLLTVM--VWV--- 216

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
Db     217 -----SSIFFFLPVFCLTVLYSLIGRKL 239


RESULT 7
US-09-077-675A-8
; Sequence 8, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  361 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-077-675A-8

  Query Match             21.9%;  Score 277;  DB 4;  Length 361;
  Best Local Similarity   31.7%;  Pred. No. 5.9e-18;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : ::| || || |:| |||| |  |||::  :|
Db      41 VTATCVALFVVGIAGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 99

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     100 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 159

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
            ::|  :   : |   : |::     ||:               |   ||:  :|:
Db     160 FVIWAVAFCSAGPIFVLVGVEHE---NGTDPWDTNECRPTEFAVRSGLLTVM--VWV--- 211

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
Db     212 -----SSIFFFLPVFCLTVLYSLIGRKL 234


RESULT 8
US-09-077-675A-13
; Sequence 13, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  366 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
```

US-09-077-675A-13

Query Match 21.9%; Score 277; DB 4; Length 366;
Best Local Similarity 31.7%; Pred. No. 6e-18;
Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;

```
Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :|||: ||:|  ||: : : ::| || || |:| ||||: |  |||::  :|
Db      46 VTATCVALFVVGIAGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 104

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :|||| ||  | |||:  |: |   ::
Db     105 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 164

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::    ||:            |   ||:  :|:
Db     165 FVIWAVAFCSAGPIFVLVGVEHE---NGTDPWDTNECRPTEFAVRSGLLTVM--VWV--- 216

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
           | :|: ||:  ::||| |:  ::
Db     217 -----SSIFFFLPVFCLTVLYSLIGRKL 239
```

RESULT 9
US-09-077-675A-3
; Sequence 3, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-3

Query Match 21.7%; Score 274; DB 4; Length 353;
Best Local Similarity 31.7%; Pred. No. 1.1e-17;
Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;

```
Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :|||: ||:|  ||: : : |:| || || |:| ||||: |  |||::: :|
Db      33 VTATCVALFVVGIAGNLLTMLVVSRFREMRTTTNLYLSSMAFSDLLIFLC-MPLDLFRLW 91

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |:  | : |     : |:  :|::|:|| :|||| ||  | |||:  |: |   ::
Db      92 QYRPWNLGNLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 151

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::    ||:            |   ||:  :|:
Db     152 LVIWAVAFCSAGPIFVLVGVE---HDNGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 203

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
           | :|: ||:  ::||| |:  ::
Db     204 -----SSVFFFLPVFCLTVLYSLIGRKL 226
```

RESULT 10
US-09-077-675A-16
; Sequence 16, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-16

Query Match 21.7%; Score 274; DB 4; Length 364;
Best Local Similarity 31.7%; Pred. No. 1.1e-17;
Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;

```
Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :|||: ||:|  ||: : : ::| || || |:| ||||: |  |||::  :|
```

```
Db      45 VTATCVALFVVGISGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 103

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     104 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 163

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::     ||:               |   ||:  :|:
Db     164 LVIWAVAFCSAGPIFVLVGVEHE---NGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 215

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
Db     216 -----SSVFFFLPVFCLTVLYSLIGRKL 238


RESULT 11
US-09-077-675A-5
; Sequence 5, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-077-675A-5

  Query Match             21.6%;  Score 273;  DB 4;  Length 289;
  Best Local Similarity   31.7%;  Pred. No. 1.1e-17;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : |:| || || |:| |:|||: |  ||||:: :|
Db      46 VTATCVALFVVGIAGNLLTMLVVSRFREMRTTTNLYLSSMAFSELLIFLC-MPLELFRLW 104

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |:  | : |     : |:  :|::|:|| :|||||  ||  | |||:  |: |   ::

Db     105 QYRPWNLGNLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 164

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::     ||:               |   ||:  :|:
Db     165 LVIWAVAFCSAGPIFVLVGVE---HDNGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 216

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
Db     217 -----SSVFFFLPVFCLTVLYSLIGRKL 239


RESULT 12
US-07-629-104I-3
; Sequence 3, Application US/07629104I
; Patent No. 5288621
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C
;    APPLICANT:  Straub, Richard E
;    TITLE OF INVENTION:  PITUITARY TRH RECEPTOR
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yahwak & Associates
;      STREET:  25 Skytop Drive
;      CITY:  Trumbull
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06611
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Microsoft Word 3.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/629,104I
;      FILING DATE:  19901214
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  George M. Yahwak
;      REGISTRATION NUMBER:  26,824
;      REFERENCE/DOCKET NUMBER:  CRF D - 995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)268-1951
;      TELEFAX:  (203)268-1951
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  393 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-629-104I-3

  Query Match             21.0%;  Score 265;  DB 1;  Length 393;
  Best Local Similarity   28.9%;  Pred. No. 8.1e-17;
  Matches   72;  Conservative   48;  Mismatches   69;  Indels   60;  Gaps    9;

Qy      15 LEDPFQKHLNSTE---------EY-------LAFLCGPRRSHFFLPVSVVYVPIFVVGVI 58
           :|:     :| ||         ||       :  :||                   :|::
Db       1 MENDTVSEMNQTELQPQAAVALEYQVVTILLVVIICG------------------LGIV 41

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---EVYEMWRNYPFLFGPV 114
           ||::| ||::: : |:|||| || ||||:||:||:  |:|     :|  |     ::| |
Db      42 GNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDSIYGSW-----VYGYV 96

Qy     115 GCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLF 174
           ||   | |       ||  |||  ::|||:|| || :|:   |  || :|:  || |: ::
Db      97 GCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWAFTSIY 156

Qy     175 SLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFIIQVTSFLFYLLPMTV 228
           :    :  :    : |  :|    :|          |::: :| :      ||::|| :
Db     157 CMLWFFLLDLNISTYKNAVVV----SCGYKISRNYYSPIYLMDFGV------FYVVPMIL 206

Qy     229 ISVLYYLMA 237
           :|||  :|
```

```
Db    207 ATVLYGFIA 215


RESULT 13
US-09-261-599B-3
; Sequence 3, Application US/09261599B
; Patent No. 6395877
; GENERAL INFORMATION:
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION:  14273 Receptor, A No. 6395877el G-Protein Coupled Receptor
;  FILE REFERENCE:  5800-4B, 035800/177086
;  CURRENT APPLICATION NUMBER: US/09/261,599B
;  CURRENT FILING DATE:  1999-02-26
;  PRIOR APPLICATION NUMBER:  09/107,761
;  PRIOR FILING DATE:  1998-06-30
;  PRIOR APPLICATION NUMBER:  09/223,538
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Seven
;   OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily
US-09-261-599B-3

  Query Match             20.8%;  Score 262.5;  DB 4;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 8.7e-17;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||||: : ::|||| :: :|||:||| ||   |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
             |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||      :|     ||:|      | :  |         |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 14
US-09-456-455A-3
; Sequence 3, Application US/09456455A
; Patent No. 6448005
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  APPLICANT:  Tsai, Fong-Ying
;  TITLE OF INVENTION: 14273 Receptor, A No. 6448005el G-Protein Coupled Receptor
;  FILE REFERENCE: MNI-204CP3
;  CURRENT APPLICATION NUMBER: US/09/456,455A
;  CURRENT FILING DATE:  1999-12-08
;  PRIOR APPLICATION NUMBER: 09/107,761
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: 09/223,538
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Seven
;   OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily
US-09-456-455A-3

  Query Match             20.8%;  Score 262.5;  DB 4;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 8.7e-17;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||||: : ::|||| :: :|||:||| ||   |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
             |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||      :|     ||:|      | :  |         |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 15
US-08-288-663A-1
; Sequence 1, Application US/08288663A
; Patent No. 5879896
;  GENERAL INFORMATION:
;    APPLICANT:  HINUMA, Shuji
;    APPLICANT:  HOSOYA, Masaki
;    APPLICANT:  ONDA, Haruo
;    TITLE OF INVENTION:  HUMAN TRH RECEPTOR, ITS PRODUCTION
;    TITLE OF INVENTION:  AND USE
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,663A
;      FILING DATE:  09-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  198309/1993
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  286986/1993
;      FILING DATE:  16-NOV-1993
;      APPLICATION NUMBER:  325215/1993
;      FILING DATE:  22-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resnick, David S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  44612
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  398 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-288-663A-1
```

```
  Query Match             20.5%;  Score 259;  DB 2;  Length 398;
  Best Local Similarity   32.4%;  Pred. No. 2.9e-16;
  Matches   66;  Conservative   47;  Mismatches   65;  Indels   26;  Gaps    7;

Qy      45 VSVVYVPIFV-VGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---E 99
           |::: | |   :|::||::| ||::: : |:|||| || ||||:||:||:  |:|
Db      27 VTILLVLIICGLGIVGNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDS 86

Qy     100 VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRR 159
           :|  |     ::| |||   | |      ||  |||  ::|||:|| || :|:   |  |
Db      87 IYGSW-----VYGYVGCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSR 141

Qy     160 ALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFI 213
           | :|:  || |: |: :    :  :    : :  ::    :|          |::: :|
Db     142 AKKIIIFVWAFTSLYCMLWFFLLDLNISTYKDAIVI----SCGYKISRNYYSPIYLMDFG 197

Qy     214 IQVTSFLFYLLPMTVISVLYYLMA 237
           :      ||::|| : :|||  :|
Db     198 V------FYVVPMILATVLYGFIA 215


Search completed: January 17, 2003, 05:08:09
Job time : 28 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:26:05 ; Search time 54 Seconds
(without alignments)
4140.136 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc.....cactcagagtgagtatctag 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	717.8	98.5	1248	4	US-09-545-944-1	Sequence 1, Appli
2	282.4	38.7	1535	4	US-09-668-680-12	Sequence 12, Appl
3	120.4	16.5	1088	4	US-09-077-675A-6	Sequence 6, Appli
4	120.4	16.5	1122	4	US-09-077-675A-9	Sequence 9, Appli
5	114	15.6	1063	4	US-09-077-675A-1	Sequence 1, Appli
6	112.4	15.4	1029	4	US-09-077-675A-4	Sequence 4, Appli
7	112.4	15.4	1161	1	US-08-086-439C-2	Sequence 2, Appli
8	112.4	15.4	1161	1	US-08-434-877-2	Sequence 2, Appli
9	112.4	15.4	1367	3	US-08-475-742-3	Sequence 3, Appli
10	112.4	15.4	1370	1	US-08-056-051-1	Sequence 1, Appli
11	112.4	15.4	1370	1	US-07-928-611-17	Sequence 17, Appl
12	112.4	15.4	1370	2	US-08-487-811A-17	Sequence 17, Appl
13	112.4	15.4	1370	4	US-09-060-694-17	Sequence 17, Appl
14	112.4	15.4	1370	4	US-09-378-074-17	Sequence 17, Appl
15	112.4	15.4	1370	5	PCT-US93-07370-17	Sequence 17, Appl
16	112.4	15.4	1466	1	US-08-056-051-3	Sequence 3, Appli
17	112.4	15.4	1466	1	US-07-928-611-19	Sequence 19, Appl
18	112.4	15.4	1466	2	US-08-487-811A-19	Sequence 19, Appl
19	112.4	15.4	1466	4	US-09-060-694-19	Sequence 19, Appl
20	112.4	15.4	1466	4	US-09-378-074-19	Sequence 19, Appl
21	112.4	15.4	1466	5	PCT-US93-07370-19	Sequence 19, Appl
22	112.4	15.4	1610	1	US-08-056-051-5	Sequence 5, Appli
23	112.4	15.4	1610	1	US-07-928-611-21	Sequence 21, Appl
24	112.4	15.4	1610	2	US-08-487-811A-21	Sequence 21, Appl
25	112.4	15.4	1610	4	US-09-060-694-21	Sequence 21, Appl
26	112.4	15.4	1610	4	US-09-378-074-21	Sequence 21, Appl
27	112.4	15.4	1610	5	PCT-US93-07370-21	Sequence 21, Appl
28	110	15.1	1529	3	US-08-858-876A-3	Sequence 3, Appli
29	110	15.1	1529	4	US-09-472-880-3	Sequence 3, Appli
30	109.2	15.0	1092	4	US-09-077-675A-15	Sequence 15, Appl
31	109.2	15.0	3129	4	US-09-077-675A-14	Sequence 14, Appl
32	99.4	13.6	1600	4	US-09-029-027B-1	Sequence 1, Appli
33	99	13.6	1233	4	US-09-200-090-1	Sequence 1, Appli
34	97.6	13.4	1575	3	US-08-858-876A-1	Sequence 1, Appli
35	97.6	13.4	1575	4	US-09-472-880-1	Sequence 1, Appli
36	97.2	13.3	283	4	US-08-993-088A-4	Sequence 4, Appli
37	97.2	13.3	283	4	US-08-993-424B-4	Sequence 4, Appli
38	95.8	13.1	1176	4	US-09-200-090-3	Sequence 3, Appli
39	93.2	12.8	1164	4	US-08-993-088A-6	Sequence 6, Appli
40	93.2	12.8	1164	4	US-08-993-424B-6	Sequence 6, Appli
41	88.8	12.2	1601	1	US-08-722-001-7	Sequence 7, Appli
42	88.8	12.2	1997	1	US-08-722-001-27	Sequence 27, Appl
43	88.8	12.2	2004	1	US-08-722-001-11	Sequence 11, Appl
44	88.6	12.2	1639	1	US-08-334-698-5	Sequence 5, Appli
45	88.6	12.2	1639	1	US-08-228-932-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-545-944-1
; Sequence 1, Application US/09545944
; Patent No. 6461836
; GENERAL INFORMATION:
; APPLICANT: AMES, ROBERT
; APPLICANT: ELSHOURBAGY, NABIL
; APPLICANT: MICHALOVICH, DAVID
; APPLICANT: SARAU, HENRY
; APPLICANT: SHABON, USMAN
; APPLICANT: VAWTER, LISA
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
; TITLE OF INVENTION: (AXOR34) AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP70657-1
; CURRENT APPLICATION NUMBER: US/09/545,944
; CURRENT FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 09/435,384
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1248
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-545-944-1

Query Match 98.5%; Score 717.8; DB 4; Length 1248;
Best Local Similarity 99.7%; Pred. No. 1.4e-149;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309
```

```
Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy    721 A 721
          |
Db    730 A 730
```

RESULT 2
US-09-668-680-12
; Sequence 12, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6436703el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2A
; CURRENT APPLICATION NUMBER: US/09/668,680
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 12
; LENGTH: 1535
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1338)
US-09-668-680-12

Query Match 38.7%; Score 282.4; DB 4; Length 1535;
Best Local Similarity 66.0%; Pred. No. 9.9e-54;
Matches 409; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

```
Qy    103 GGACCTCGGCGCAGCCACTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTG 162
          || || | ||  |   |  | ||| | ||| | | ||     ||  |||  || || |||

Db    214 GGGCCCCAGCAGACAGAGCTGTTCATGCCCATCTGTGCCACATACCTGCTGATCTTCGTG 273

Qy    163 GTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAG 222
          ||||| |   | |||||||  |||   || ||||| || ||||  ||| |||| |||

Db    274 GTGGGCGCTGTGGGCAATGGGCTGACCTGTCTGGTCATCCTGCGCCACAAGGCCATGCGC 333

Qy    223 ACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTT 282
          ||||| ||||||||||||||||||||||||| || || ||||| ||||| |||||  |

Db    334 ACGCCTACCAACTACTACCTCTTCAGCCTGGCCGTGTCGGACCTGCTGGTGCTGCTGGTG 393

Qy    283 GGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTG 342
          ||  |||||||||| |||||||||||||| |||||||| ||| || | ||    |

Db    394 GGCCTGCCCCTGGAGCTCTATGAGATGTGGCACAACTACCCCTTCCTGCTGGGCGTTGGT 453

Qy    343 GGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATC 402
          |||||||| |||   |||   || |||||||  || ||| | |||||  | |||| | ||

Db    454 GGCTGCTATTTCCGCACGCTACTGTTTGAGATGGTCTGCCTGGCCTCAGTGCTCAACGTC 513

Qy    403 ACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAG 462
          ||  || | |||||||| ||||| ||||||  |  | |||||  |||  ||||       |

Db    514 ACTGCCCTGAGCGTGGAACGCTATGTGGCCGTGGTGCACCCACTCCAGGCCAGGTCCATG 573

Qy    463 AGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTC 522
             || |||  ||  |  | | |  | || ||   |||||||||  |  || ||||||  |

Db    574 GTGACGCGGGCCCATGTGCGCCGAGTGCTTGGGGCCGTCTGGGGTCTTGCCATGCTCTGC 633

Qy    523 TCCCTGCCCAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTG 582
          |||||||||||||||||  | || ||||||  |   |   || | |||    ||  ||

Db    634 TCCCTGCCCAACACCAGCCTGCACGGCATCCGGCAGCTGCACGTGCCCTGCCGGGGCCCA 693

Qy    583 GTCCCAGGTTCGGCCACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATC 642
          || ||||  || ||    || | | |  ||  |||   |    |||||||  |  |  |

Db    694 GTGCCAGACTCAGCTGTTTGCATGCTGGTCCCGCCACGGGCCCTCTACAACATGGTAGTG 753

Qy    643 CAGGTCACCTCCTTCCTATTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTAC 702
          |||  |||| |  | || |||| |  ||| ||||||  |  |||| || || ||||||

Db    754 CAGACCACCGCGCTGCTCTTCTTCTGCCTGCCCATGGCCATCATGAGCGTGCTCTACCTG 813

Qy    703 CTCATGGCACTCAGAGTGAG 722
          ||||| |  ||  || || |
Db    814 CTCATTGGGCTGCGACTGCG 833
```

RESULT 3
US-09-077-675A-6
; Sequence 6, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1088 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-077-675A-6

  Query Match             16.5%;  Score 120.4;  DB 4;  Length 1088;
  Best Local Similarity   52.3%;  Pred. No. 4.2e-18;
  Matches  290;  Conservative    0;  Mismatches  261;  Indels    3;  Gaps    1;

Qy     132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
           |||  | |      |  ||| || | || ||||||||  ||  ||||||  | ||
Db     122 CGTCACAGCCACCTGCGTGGCACTCTTCGTGGTGGGTATCGCTGGCAACCTGCTCACCAT 181

Qy     192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
           ||||||  |   ||    ||  |   ||   ||  ||||||||  ||||||  | |||| |
Db     182 GCTGGTGGTGTCGCGCTTCCGCGAGCTGCGCACCACCACCAACCTCTACCTGTCCAGCAT 241

Qy     252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
           |||  |||| ||    ||| || |  ||||  | ||||||||||  ||  |    | ||
Db     242 GGCCTTCTCCGAT---CTGCTCATCTTCCTCTGCATGCCCCTGGACCTCGTTCGCCTCTG 298

Qy     312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
           ||   ||   || |     |||||   | |   |||| |  ||         | ||  |||
Db     299 GCAGTACCGGCCCTGGAACTTCGGCGACCTCCTCTGCAAACTCTTCCAATTCGTCAGTGA 358

Qy     372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
           || |     || |||| |  | |||| ||||||  |  | |||||  |||||||||  | ||
Db     359 GAGCTGCACCTACGCCACGGTGCTCACCATCACAGCGCTGAGCGTCGAGCGCTACTTCGC 418

Qy     432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
           ||||     |||  |||| |||||  ||  |  ||||  | | ||||        | || |
Db     419 CATCTGCTTCCCACTCCGGGCCAAGGTGGTGGTCACCAAGGGGCGGGTGAAGCTGGTCAT 478

Qy     492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCAT 551
           |  | || ||||||  | |  || |   |  | ||  |||||  |  |    |    ||  |
Db     479 CTTCGTCATCTGGGCCGTGGCCTTCTGCAGCGCCGGGCCCATCTTCGTGCTAGTCGGGGT 538

Qy     552 CAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGGTCAT 611
             ||  | |  ||  | || |    |       |   |  |    | ||     | |
Db     539 GGAGCACGAGAACGGCACCGACCCTTGGGACACCAACGAGTGCCGCCCCACCGAGTTTGC 598

Qy     612 CAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACCTCCT 671
             || |    |  |   ||   ||||  |   |||  ||  | || | |||| ||| |
Db     599 GGTGCGCTCTGGACTGCTCACGGTCATGGTGTGGGTGTCCAGCATCTTCTTCTTCCTTCC 658

Qy     672 CCCCATGACTGTCA 685
           |  |   | |||
Db     659 TGTCTTCTGTCTCA 672


RESULT 4
US-09-077-675A-9
; Sequence 9, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-077-675A-9

  Query Match             16.5%;  Score 120.4;  DB 4;  Length 1122;
  Best Local Similarity   52.3%;  Pred. No. 4.2e-18;
  Matches  290;  Conservative    0;  Mismatches  261;  Indels    3;  Gaps    1;

Qy     132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
           |||  | |      |  ||| || | || ||||||||  ||  ||||||  | ||
Db     387 CGTCACAGCCACCTGCGTGGCACTCTTCGTGGTGGGTATCGCTGGCAACCTGCTCACCAT 446

Qy     192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
           ||||||  |   ||    ||  |   ||   ||  ||||||||  ||||||  | |||| |
Db     447 GCTGGTGGTGTCGCGCTTCCGCGAGCTGCGCACCACCACCAACCTCTACCTGTCCAGCAT 506

Qy     252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
           |||  |||| ||    ||| || |  ||||  | ||||||||||  ||  |    | ||
Db     507 GGCCTTCTCCGAT---CTGCTCATCTTCCTCTGCATGCCCCTGGACCTCGTTCGCCTCTG 563

Qy     312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
           ||   ||   || |     |||||   | |   |||| |  ||         | ||  |||
Db     564 GCAGTACCGGCCCTGGAACTTCGGCGACCTCCTCTGCAAACTCTTCCAATTCGTCAGTGA 623

Qy     372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
           || |     || |||| |  | |||| ||||||  |  | |||||  |||||||||  | ||
Db     624 GAGCTGCACCTACGCCACGGTGCTCACCATCACAGCGCTGAGCGTCGAGCGCTACTTCGC 683

Qy     432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
           ||||     |||  |||| |||||  ||  |  ||||  | | ||||        | || |
Db     684 CATCTGCTTCCCACTCCGGGCCAAGGTGGTGGTCACCAAGGGGCGGGTGAAGCTGGTCAT 743

Qy     492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCAT 551
           |  | || ||||||  | |  || |   |  | ||  |||||  |  |    |    ||  |
Db     744 CTTCGTCATCTGGGCCGTGGCCTTCTGCAGCGCCGGGCCCATCTTCGTGCTAGTCGGGGT 803

Qy     552 CAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGGTCAT 611
             ||  | |  ||  | || |    |       |   |  |    | ||     | |
Db     804 GGAGCACGAGAACGGCACCGACCCTTGGGACACCAACGAGTGCCGCCCCACCGAGTTTGC 863
```

```
Qy    612 CAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACCTCCT 671
          || |    |  |   ||   |||| |   |||  ||  | || | |||| ||| |
Db    864 GGTGCGCTCTGGACTGCTCACGGTCATGGTGTGGGTGTCCAGCATCTTCTTCTTCCTTCC 923

Qy    672 CCCCATGACTGTCA 685
          | |   | |||
Db    924 TGTCTTCTGTCTCA 937


RESULT 5
US-09-077-675A-1
; Sequence 1, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
     TITLE OF INVENTION:  RECEPTOR ASSAY
     NUMBER OF SEQUENCES:  16
     CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
       TELEX:
   INFORMATION FOR SEQ ID NO:  1:
     SEQUENCE CHARACTERISTICS:
,      LENGTH:  1063 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-077-675A-1

  Query Match           15.6%;  Score 114;  DB 4;  Length 1063;
  Best Local Similarity  52.2%;  Pred. No. 1.1e-16;
  Matches  302;  Conservative    0;  Mismatches  270;  Indels    6;  Gaps    2;

Qy    132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
          |||  | |     |  ||| |  | || ||||||||  ||   |||||  | ||   |
Db     97 CGTCACCGCCACCTGCGTGGCGCTCTTCGTGGTGGGTATCGCGGGCAACCTGCTCACGAT 156

Qy    192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
          |||||  |    |    ||  |  |||   ||  |||||||| ||||||  | |||| |
Db    157 GCTGGTAGTGTCACGCTTCCGCGAGATGCGCACCACCACCAACCTCTACCTGTCCAGCAT 216

Qy    252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
          |||  |||| |||||  |   || |  ||||  | ||||||||  ||  |||      | ||
Db    217 GGCCTTCTCCGACCTAC---TCATCTTCCTCTGCATGCCCCTCGACCTCTTCCGCCTCTG 273

Qy    312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
          ||   ||   ||||       | ||   | ||  |||| |  ||     |  | |    ||
Db    274 GCAGTACCGGCCTTGGAACCTTGGCAACCTGCTCTGCAAACTCTTCCAGTTCGTTAGCGA 333

Qy    372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
          || |      || |||| |  | |||| |||||||  |  | |||||  |||||||| | ||
Db    334 GAGCTGCACCTACGCCACAGTGCTCACCATCACCGCGCTGAGCGTCGAGCGCTACTTCGC 393

Qy    432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
          ||||      |||| | || |||||  |   |  ||||  | ||||||       | || |
Db    394 CATCTGCTTCCCGCTGCGGGCCAAGGTAGTGGTCACCAAGGGCCGGGTAAAGCTGGTCAT 453

Qy    492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCAT 551
          |    || |||||| | |  || |   |  | ||  |||| |  |    |    || |
Db    454 CCTGGTCATCTGGGCCGTGGCCTTCTGCAGCGCCGGGCCCATCTTCGTGCTGGTCGGAGT 513

Qy    552 CAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGGTCAT 611
            ||    |  ||  | |  |             |   |  |    | ||     | ||
Db    514 GGAGCATGATAACGGCACTGACCCTCGGGACACCAACGAGTGCCGCGCCACGGAGTTCGC 573

Qy    612 CAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACCTCCT 671
          |  || |    || |     |   |||| ||  |||  ||    || | |||   ||||
Db    574 CGTGCGCTCCGGGCTGCTTACCGTCATGGTCTGGGTGTCCAGTGTCTTCTTCT---TCCT 630

Qy    672 CCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGG 709
          ||  |      | ||| ||| |||||   |||||| |
Db    631 GCCTGTCTTCTGCCTCACTGTGCTCTATAGCCTCATCG 668


RESULT 6
US-09-077-675A-4
; Sequence 4, Application US/09077675A
; Patent No. 6242199
;  GENERAL INFORMATION:
;    APPLICANT:  Pai, Lee-Yuh
;    APPLICANT:  Feighner, Scott C.
;    APPLICANT:  Howard, Andrew D.
;    APPLICANT:  Pong, Sheng-Shung
;    APPLICANT:  Van Der Ploeg, Leonardus H.T.
;    TITLE OF INVENTION:  RECEPTOR ASSAY
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/077,675A
;      FILING DATE:  3-JUN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cocuzzo, Anna L.
;      REGISTRATION NUMBER:  42,452
;      REFERENCE/DOCKET NUMBER:  19590P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-1273
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1029 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-077-675A-4

  Query Match             15.4%;  Score 112.4;  DB 4;  Length 1029;
  Best Local Similarity   52.1%;  Pred. No. 2.4e-16;
  Matches  301;  Conservative    0;  Mismatches  271;  Indels    6;  Gaps    2;

Qy    132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
          ||| | |     |  ||| |  | || ||||||||  ||   |||||  | ||   |
Db    294 CGTCACCGCCACCTGCGTGGCGCTCTTCGTGGTGGGTATCGCGGGCAACCTGCTCACGAT 353

Qy    192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
          |||||  |    |    ||  |  |||   ||  ||||||||  ||||||  | |||| |
Db    354 GCTGGTAGTGTCACGCTTCCGCGAGATGCGCACCACCACCAACCTCTACCTGTCCAGCAT 413

Qy    252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
          |||  |||| || || |   || |  ||||  | ||||||| ||  |||       | ||
Db    414 GGCCTTCTCCGAACTAC---TCATCTTCCTCTGCATGCCCCTCGAACTCTTCCGCCTTTG 470

Qy    312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
          ||   ||   ||||      | ||   | ||  |||| |  ||     |  | |    ||
Db    471 GCAGTACCGGCCTTGGAACCTTGGCAACCTGCTCTGCAAACTCTTCCAGTTCGTTAGCGA 530

Qy    372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
          || |     || |||| |  | |||| |||||||  |  | |||||  |||||||| | ||
Db    531 GAGCTGCACCTACGCCACAGTGCTCACCATCACCGCGCTGAGCGTCGAGCGCTACTTCGC 590

Qy    432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
          ||||     |||| | || |||||  |   |  ||||  | ||||||     | || |
Db    591 CATCTGCTTCCCGCTGCGGGCCAAGGTAGTGGTCACCAAGGGCCGGGTAAAGCTGGTCAT 650

Qy    492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCAT 551
          |    || |||||| | |  || |   | | ||  |||| |  |     |    ||  |
Db    651 CCTGGTCATCTGGGCCGTGGCCTTCTGCAGCGCCGGGCCCATCTTCGTGCTGGTCGGAGT 710

Qy    552 CAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGGTCAT 611
            ||    |  ||  | |  |            |   |  |    | ||     | ||
Db    711 GGAGCATGATAACGGCACTGACCCTCGGGACACCAACGAGTGCCGCGCCACGGAGTTCGC 770

Qy    612 CAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACCTCCT 671
          |  || |    || |    |  |||| ||  |||  ||    || | ||||   ||||
Db    771 CGTGCGCTCCGGGCTGCTTACCGTCATGGTCTGGGTGTCCAGTGTCTTCTTCT---TCCT 827

Qy    672 CCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGG 709
            ||  |     | ||| ||| |||||   |||||| |
Db    828 GCCTGTCTTCTGCCTCACTGTGCTCTATAGCCTCATCG 865


RESULT 7
US-08-086-439C-2
; Sequence 2, Application US/08086439C
; Patent No. 5468615
;  GENERAL INFORMATION:
;    APPLICANT:  Chio, Christopher L.
;    APPLICANT:  Huff, Rita M.
;    TITLE OF INVENTION:  A Synthetic Gene for D4 Dopamine
;    TITLE OF INVENTION:  Receptors
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Upjohn Company, Corp. Intellectual
;      ADDRESSEE:  Property Law
;      STREET:  301 Henrietta Street
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/086,439C
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Darnley Jr., James D.
;      REGISTRATION NUMBER:  33,673
;      REFERENCE/DOCKET NUMBER:  4700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  616-385-5210
;      TELEFAX:  616-385-6897
;      TELEX:  224401
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1161 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-086-439C-2

  Query Match             15.4%;  Score 112.4;  DB 1;  Length 1161;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy    139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
          |||| |   ||||    |  | |  | || | ||   || ||    || ||||||  ||
Db    112 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 171

Qy    199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
          |         | |  ||  || ||||||||||||||| || | || | ||||||||| |
Db    172 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 231

Qy    259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
          | ||||||||  ||   |||||  |   |||| ||    |||||      |   |
Db    232 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 291

Qy    319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
          | | |   || |  | |||       ||| ||  |   | |||| |    |  |   ||
Db    292 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 351

Qy    379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
          |||  |||||||||| ||| | |   | || ||||||||||  | | |||||| |
Db    352 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 411

Qy    439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
          ||| | |||  ||| | |||| |      | ||||||   ||   | || |||||  |
Db    412 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 471

Qy    499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
          |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db    472 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 531

Qy    559 CACTACTTCCCCAATGGGTCCC 580
            |  |  ||||   | || ||
Db    532 GGCCGCGACCCCGCCGTGTGCC 553


RESULT 8
US-08-434-877-2
; Sequence 2, Application US/08434877
; Patent No. 5721132
;  GENERAL INFORMATION:
;    APPLICANT:  Chio, Christopher L.
;    APPLICANT:  Huff, Rita M.
;    TITLE OF INVENTION:  A Synthetic Gene for D4 Dopamine
;    TITLE OF INVENTION:  Receptors
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Upjohn Company, Corp. Intellectual
;      ADDRESSEE:  Property Law
;      STREET:  301 Henrietta Street
;      CITY:  Kalamazoo
```

```
;       STATE:  Michigan
;       COUNTRY:  USA
;       ZIP:  49001
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette (DS,HD)
;       COMPUTER:  Gateway 2000, P5-90
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/434,877
;       FILING DATE:  1 July 1993
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Darnley Jr., James D.
;       REGISTRATION NUMBER:  33,673
;       REFERENCE/DOCKET NUMBER:  4700 DV1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  616-385-5210
;       TELEFAX:  616-385-6897
;       TELEX:  224401
    INFORMATION FOR SEQ ID NO:  2:
      SEQUENCE CHARACTERISTICS:
        LENGTH:  1161 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-434-877-2

  Query Match             15.4%;  Score 112.4;  DB 1;  Length 1161;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||   | | | | || | ||   || ||    || ||||||  ||
Db     112 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 171

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
           |       | |  ||  || |||||||||||||||| || | || | ||||||||| |
Db     172 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 231

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | ||||||||  ||   ||||| |   |||| ||    |||||     |   |
Db     232 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 291

Qy     319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
           | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     292 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 351

       379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  |||||||||| ||| | |  | || ||||||||||  | | |||||| |
Db     352 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 411

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
           |||  | |||  ||| | |||| |     | ||||||   ||   | || |||||  |
Db     412 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 471

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
           |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db     472 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 531

Qy     559 CACTACTTCCCCAATGGGTCCC 580
           |  |  ||||   | || ||
Db     532 GGCCGCGACCCCGCCGTGTGCC 553


RESULT 9
US-08-475-742-3
; Sequence 3, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
;  APPLICANT: O'Malley, Karen L
;  APPLICANT: Todd, Richard D
;  TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
;  FILE REFERENCE: WU 102 CON DIV
;  CURRENT APPLICATION NUMBER: US/08/475,742
;  CURRENT FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: US 08/261,293
;  EARLIER FILING DATE: 1994-06-16
;  EARLIER APPLICATION NUMBER: US 08/014,013
;  EARLIER FILING DATE: 1993-01-28
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1367
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1367)
;   OTHER INFORMATION: D4 Dopamine Receptor cDNA
;  PUBLICATION INFORMATION:
;   AUTHORS: Van tol, H. H.
;   AUTHORS: Bunzow, J. R.
;   TITLE: Cloning of the gene for a human dopamine D4 receptor
;   TITLE: with high affinity for the antipsychotic clozapine
;   JOURNAL: Nature
;   VOLUME: 350
;   PAGES: 610-614
;   DATE: 1991
US-08-475-742-3

  Query Match             15.4%;  Score 112.4;  DB 3;  Length 1367;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||   | | | | || | ||   || ||    || ||||||  ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
           |       | |  ||  || |||||||||||||||| || | || | ||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | ||||||||  ||   ||||| |   |||| ||    |||||     |   |
Db     338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy     319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
           | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy     379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  |||||||||| ||| | |  | || ||||||||||  | | |||||| |
Db     458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
           |||  | |||  ||| | |||| |     | ||||||   ||   | || |||||  |
Db     518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
           |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db     578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy     559 CACTACTTCCCCAATGGGTCCC 580
           |  |  ||||   | || ||
Db     638 GGCCGCGACCCCGCCGTGTGCC 659


RESULT 10
US-08-056-051-1
; Sequence 1, Application US/08056051
; Patent No. 5516683
;  GENERAL INFORMATION:
;    APPLICANT:  Grandy, David K
;    APPLICANT:  Bunzow, James R
```

```
;    APPLICANT:  Civelli, Olivier
;    APPLICANT:  Van Tol, Hubert H.-M.
;    TITLE OF INVENTION:  A No. 5516683el Human Dopamine Receptor and Uses
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/056,051
       FILING DATE:  19930429
       CLASSIFICATION:  435
     ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5516683nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  90,1092-C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..103
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  1268..1370
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  104..1267
US-08-056-051-1

  Query Match             15.4%;  Score 112.4;  DB 1;  Length 1370;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||   |  | |  | || | ||   || ||    || |||||| ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
             |         | |  ||  || |||||||||||||| || | || | |||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | ||||||||  ||   ||||| |   |||| ||    |||||      |   |
Db     338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy     319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
             | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy     379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  ||||||||||| ||| | |   | || ||||||||||  | | |||||||  |
Db     458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
             ||| | |||  ||| | |||| |     | ||||||   ||   | || |||||  |
Db     518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
           |||   | |||| |   |  |   ||||      ||  ||    |  |  |  |
Db     578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy     559 CACTACTTCCCCAATGGGTCCC 580
             |  |  ||||   | || ||
Db     638 GGCCGCGACCCCGCCGTGTGCC 659


RESULT 11
US-07-928-611-17
; Sequence 17, Application US/07928611
; Patent No. 5569601
;  GENERAL INFORMATION:
;    APPLICANT:  Van Tol, Hubert H.M.
;    APPLICANT:  Civelli, Olivier
;    TITLE OF INVENTION:  A No. 5569601el Human Dopamine Receptor and Uses
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/928,611
;      FILING DATE:  19920810
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5569601nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  90,1092-B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  810-221-8317
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..103
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  1268..1370
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  104..1267
US-07-928-611-17

  Query Match             15.4%;  Score 112.4;  DB 1;  Length 1370;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||   |  | |  | || | ||   || ||    || |||||| ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
             |         | |  ||  || |||||||||||||| || | || | |||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337
```

```
Qy    259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
Db    338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy    319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
Db    398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy    379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
Db    458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy    439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
Db    518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy    499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
Db    578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy    559 CACTACTTCCCCAATGGGTCCC 580
Db    638 GGCCGCGACCCCGCCGTGTGCC 659
```

RESULT 12
US-08-487-811A-17
; Sequence 17, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..103
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1268..1370
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 104..1267
US-08-487-811A-17

Query Match 15.4%; Score 112.4; DB 2; Length 1370;
Best Local Similarity 53.4%; Pred. No. 2.5e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```
Qy    139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
Db    218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy    199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
Db    278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy    259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
Db    338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy    319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
Db    398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy    379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
Db    458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy    439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
Db    518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy    499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
Db    578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy    559 CACTACTTCCCCAATGGGTCCC 580
Db    638 GGCCGCGACCCCGCCGTGTGCC 659
```

RESULT 13
US-09-060-694-17
; Sequence 17, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:

```
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..103
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  1268..1370
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  104..1267
US-09-060-694-17

 Query Match             15.4%;  Score 112.4;  DB 4;  Length 1370;
 Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
 Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||    |  | |  | || | ||   || ||    || |||||| ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
             |      | |  ||  || ||||||||||||||| || | || | |||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | ||||||||  ||   ||||| |   |||| ||    |||||      |   |
Db     338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy     319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
             | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy     379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  |||||||||| ||| | |   | || ||||||||||  | | |||||||  |
Db     458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
             ||| | |||  ||| |  |||| |     | ||||||   ||   | || |||||  |
Db     518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
             |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db     578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy     559 CACTACTTCCCCAATGGGTCCC 580
             |  |  ||||   | || ||
Db     638 GGCCGCGACCCCGCCGTGTGCC 659


RESULT 14
US-09-378-074-17
; Sequence 17, Application US/09378074
; Patent No. 6437114
;   GENERAL INFORMATION:
;        APPLICANT: Van Tol, Hubert H.M.
;                   Civelli, Olivier
;        TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
;        NUMBER OF SEQUENCES: 22
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Allegretti & Witcoff, Ltd.
;             STREET: 10 South Wacker Drive, Suite 3000
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: USA
;             ZIP: 60606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/378,074
;             FILING DATE: 20-Aug-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 07/928,611
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: No. 6437114nan, Kevin E
;             REGISTRATION NUMBER: 35,303
;             REFERENCE/DOCKET NUMBER: 90,1092-B
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312-715-1000
;             TELEFAX: 312-715-1234
;             TELEX: 810-221-8317
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1370 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  5'UTR
;             LOCATION:  1..103
;        FEATURE:
;             NAME/KEY:  3'UTR
;             LOCATION:  1268..1370
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  104..1267
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-378-074-17

  Query Match             15.4%;  Score 112.4;  DB 4;  Length 1370;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||    |  | |  | || | ||   || ||    || |||||| ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
             |      | |  ||  || ||||||||||||||| || | || | |||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | ||||||||  ||   ||||| |   |||| ||    |||||      |   |
Db     338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

Qy     319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
             | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy     379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  |||||||||| ||| | |   | || ||||||||||  | | |||||||  |
Db     458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
             ||| | |||  ||| |  |||| |     | ||||||   ||   | || |||||  |
Db     518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
             |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db     578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy     559 CACTACTTCCCCAATGGGTCCC 580
             |  |  ||||   | || ||
Db     638 GGCCGCGACCCCGCCGTGTGCC 659
```

```
RESULT 15
PCT-US93-07370-17
; Sequence 17, Application PC/TUS9307370
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A Novel Human Dopamine Receptor and Uses
;    NUMBER OF SEQUENCES:  22
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07370
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
     FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..103
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  1268..1370
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  104..1267
PCT-US93-07370-17

  Query Match             15.4%;  Score 112.4;  DB 5;  Length 1370;
  Best Local Similarity   53.4%;  Pred. No. 2.5e-16;
  Matches  236;  Conservative    0;  Mismatches  206;  Indels    0;  Gaps    0;

Qy     139 GTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTG 198
           |||| |   ||||   |  | |  | || | ||   || ||    || |||||| ||
Db     218 GTGGGGGGCGTGCTGCTCATCGGCGCGGTGCTCGCGGGGAACTCGCTCGTGTGCGTGAGC 277

Qy     199 ATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTC 258
             |       | |  ||  || ||||||||||||||| || | || | |||||||||| |
Db     278 GTGGCCACCGAGCGCGCCCTGCAGACGCCCACCAACTCCTTCATCGTGAGCCTGGCGGCC 337

Qy     259 TCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAAC 318
           | |||||||  ||   ||||| |   |||| ||   |||||      |   |
Db     338 GCCGACCTCCTCCTCGCTCTCCTGGTGCTGCCGCTCTTCGTCTACTCCGAGGTCCAGGGT 397

       319 TACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTG 378
           | | |   || |  | |||      ||| ||  |   | |||| |    |  |   ||
Db     398 GGCGCGTGGCTGCTGAGCCCCCGCCTGTGCGACGCCCTCATGGCCATGGACGTCATGCTG 457

Qy     379 TGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTA 438
           |||  |||||||||| ||| | |   | || |||||||||  | | |||||| |
Db     458 TGCACCGCCTCCATCTTCAACCTGTGCGCCATCAGCGTGGACAGGTTCGTGGCCGTGGCC 517

Qy     439 CACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATC 498
             ||| | |||  ||| | |||| |     | ||||||   ||   | || |||||  |
Db     518 GTGCCGCTGCGCTACAACCGGCAGGGTGGGAGCCGCCGGCAGCTGCTGCTCATCGGCGCC 577

Qy     499 GTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTC 558
             |||    | |||| |    |  |   ||||       ||  ||    |  |  |  |
Db     578 ACGTGGCTGCTGTCCGCGGCGGTGGCGGCGCCCGTACTGTGCGGCCTCAACGACGTGCGC 637

Qy     559 CACTACTTCCCCAATGGGTCCC 580
           |  |  ||||   | || ||
Db     638 GGCCGCGACCCCGCCGTGTGCC 659


Search completed: January 17, 2003, 02:35:14
Job time : 61 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:36:41 ; Search time 72 Seconds
(without alignments)
447.870 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1263	100.0	242	22	AAB68333	Amino acid sequenc
2	1252	99.1	249	22	AAG80937	Human nGPCR15. Ho
3	1252	99.1	293	22	AAB67807	Splice variant of
4	1252	99.1	296	22	AAB67806	Splice variant of
5	1252	99.1	412	22	AAB67803	Amino acid sequenc
6	1252	99.1	412	22	AAB67805	Amino acid sequenc
7	1252	99.1	415	21	AAY71296	Human orphan G pro
8	1252	99.1	415	21	AAB02830	Human G protein co
9	1252	99.1	415	21	AAY52992	Human neurotensin-
10	1252	99.1	415	22	AAG63353	Amino acid sequenc
11	1252	99.1	415	22	AAG63366	Amino acid sequenc
12	1252	99.1	415	22	AAG64297	Human GTP-binding
13	1252	99.1	415	22	AAE03629	Human G-protein co
14	1252	99.1	415	22	AAB67802	Amino acid sequenc
15	1252	99.1	415	22	AAB67804	Amino acid sequenc
16	1252	99.1	415	23	AAU77155	Human G-protein co
17	1252	99.1	415	23	AAE14262	Human NMUR2 protei
18	1033.5	81.8	395	22	AAG63367	Amino acid sequenc
19	1033.5	81.8	395	22	AAE03634	Rat G-protein coup
20	1033.5	81.8	395	23	AAE14263	Rat NMUR2 protein.
21	680.5	53.9	403	21	AAY90638	Human G protein-co
22	680.5	53.9	403	21	AAY90673	Human mutant G pro
23	680.5	53.9	403	21	AAY44642	Human growth hormo
24	680.5	53.9	403	22	AAB99185	Human FM-3. Homo
25	680.5	53.9	426	22	AAE03628	Human G-protein co
26	677.5	53.6	445	22	AAU68523	Human novel cytoki
27	677.5	53.6	445	22	AAU68566	Human novel cytoki
28	671.5	53.2	426	23	AAE15631	Human G-protein co
29	646	51.1	413	22	AAE03635	Rat G-protein coup
30	642	50.8	412	22	AAB99199	Rat FM-3. Rattus
31	642	50.8	439	22	AAE03636	Rat G-protein coup
32	596	47.2	405	21	AAY44641	Mouse growth hormo
33	596	47.2	405	22	AAB99186	Murine FM-3. Mus
34	418	33.1	419	22	AAB86964	D. melanogaster pe
35	418	33.1	428	22	ABB64752	Drosophila melanog
36	418	33.1	428	22	AAU38960	Drosophila G-prote
37	352	27.9	595	22	ABB71697	Drosophila melanog
38	352	27.9	595	22	AAU38987	Drosophila G-prote
39	337	26.7	660	22	ABB71689	Drosophila melanog
40	337	26.7	660	22	AAU38986	Drosophila G-prote
41	333	26.4	386	21	AAY54146	Amino acid sequenc
42	333	26.4	386	22	AAB62653	Short form of moti
43	333	26.4	412	21	AAB02854	Human G protein co
44	333	26.4	412	21	AAY54145	Amino acid sequenc
45	333	26.4	412	22	AAB62652	Long form of motil

ALIGNMENTS

RESULT 1
AAB68333
ID AAB68333 standard; Protein; 242 AA.
XX
AC AAB68333;
XX
DT 09-JUL-2001 (first entry)
XX
DE Amino acid sequence of human g-protein coupled receptor PFI-002.
XX
KW G-protein coupled receptor; obesity; signal transduction; diabetes;
KW metabolic disease; neurological disease; psychotherapy; dermatology;
KW urogenital disease; inflammation; cancer; tissue repair; photoageing;
KW skin pigmentation; frailty; osteoporosis; cardiovascular disease;
KW gastrointestinal disease; infection; allergy; respiratory disease;
KW sensory organ disorder; sleep disorder; hair loss; gene therapy;
KW PFI-002.
XX
OS Homo sapiens.
XX
PN EP1090990-A1.
XX
PD 11-APR-2001.
XX
PF 06-OCT-2000; 2000EP-0308852.
XX
PR 08-OCT-1999; 99GB-0023888.
XX
PA (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
PI Harland L;
XX

DR WPI; 2001-302046/32.
DR N-PSDB; AAF85107.
XX
PT New human G-protein coupled receptor (GPCR) polynucleotides and
PT polypeptides, for screening modulators of the polypeptide useful in
PT treating diseases associated with signal transduction, e.g. cancer,
PT inflammation, or especially, obesity -
XX
PS Claim 22; Page 42-43; 53pp; English.
XX
CC The present sequence represents a human G-protein coupled receptor. The
CC G-protein coupled receptor polynucleotide and polypeptide are useful
CC as pharmaceuticals or in the manufacture of medicaments for the
CC treatment of obesity. They are useful in the diagnosis and treatment
CC of diseases and disorders associated with signal transduction such as
CC obesity, diabetes and metabolic disease, neurological disease,
CC psychotherapeutics, urogenital disease, inflammation, cancer, tissue
CC repair, dermatology, skin pigmentation, photoageing, frailty,
CC osteoporosis, cardiovascular disease, gastrointestinal disease,
CC infection, allergy and respiratory disease, sensory organ disorders,
CC sleep disorders and hair loss. The polynucleotide may also by useful
CC in gene therapy.
XX
SQ Sequence 242 AA;

Query Match 100.0%; Score 1263; DB 22; Length 242;
Best Local Similarity 100.0%; Pred. No. 1e-142;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180

Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240

Qy    241 SI 242
          ||
Db    241 SI 242
```

RESULT 2
AAG80937
ID AAG80937 standard; Protein; 249 AA.
XX
AC AAG80937;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR15.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX
PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
DR WPI; 2001-389826/41.
DR N-PSDB; AAH50977.
XX
PT New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
PS Claim 37; Page 79; 261pp; English.
XX
CC The present invention relates to novel G protein-coupled receptors
CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC sequence is one such G protein-coupled receptor. GPCRs are also known as
CC seven transmembrane receptors and function in signal transduction. The
CC nGPCRx coding sequences are useful for screening a human to diagnose a
CC disorder affecting the brain or a genetic predisposition, specifically
CC schizophrenia. nGPCRx are useful for identifying compounds useful for
CC treating schizophrenia. Detection of nGPCRx in a sample is useful as a
CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC metabolic and cardiovascular diseases, proliferative disorders and
CC hormonal disorders. Modulators of nGPCRx activity have the utility for
CC treating neurological disorders, including schizophrenia, ADHD/ADD
CC (attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC migraine and senile dementia. Additional disorders include inflammatory
CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC diseases e.g. inflammatory bowel disease.
XX
SQ Sequence 249 AA;

Query Match 99.1%; Score 1252; DB 22; Length 249;
Best Local Similarity 99.2%; Pred. No. 2.3e-141;
Matches 240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 62

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db     63 VLVCLVILQHQAMKTPNTYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 122

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 182
```

```
Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 242

Qy    241 SI 242
          ||
Db    243 SI 244
```

RESULT 3
AAB67807
ID AAB67807 standard; Protein; 293 AA.
XX
AC AAB67807;
XX
DT 29-JUN-2001 (first entry)
XX
DE Splice variant of G-protein coupled receptor IGS4A short version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAF80327.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Example 1b; Page 96-97; 102pp; English.
XX
CC The present sequence represents a splice variant of the short version of
CC a human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
CC polynucleotides are useful for preventing, ameliorating or correcting
CC dysfunctions or diseases. These diseases include peripheral nervous
CC system, psychiatric and central nervous system disorders
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC (e.g. heart failure, angina pectoris, myocardial infarction or
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC disorders (e.g. inflammatory bowel disease or motility disorders),
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 293 AA;

```
  Query Match            99.1%;  Score 1252;  DB 22;  Length 293;
  Best Local Similarity  99.6%;  Pred. No. 2.8e-141;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180

Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 240
```

RESULT 4
AAB67806
ID AAB67806 standard; Protein; 296 AA.
XX
AC AAB67806;
XX
DT 29-JUN-2001 (first entry)
XX
DE Splice variant of G-protein coupled receptor IGS4A long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAB80326.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Example 1b; Page 92-93; 102pp; English.
XX

CC The present sequence represents a splice variant of the long version of a
CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
CC polynucleotides are useful for preventing, ameliorating or correcting
CC dysfunctions or diseases. These diseases include peripheral nervous
CC system, psychiatric and central nervous system disorders
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC (e.g. heart failure, angina pectoris, myocardial infarction or
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC disorders (e.g. inflammatory bowel disease or motility disorders),
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 296 AA;

Query Match 99.1%; Score 1252; DB 22; Length 296;
Best Local Similarity 99.6%; Pred. No. 2.8e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243
```

RESULT 5
AAB67803
ID AAB67803 standard; Protein; 412 AA.
XX
AC AAB67803;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of G-protein coupled receptor IGS4A short version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAF80323.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 18; Page 81-82; 102pp; English.
XX
CC The present sequence represents the short version of a human G-protein
CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 412 AA;

Query Match 99.1%; Score 1252; DB 22; Length 412;
Best Local Similarity 99.6%; Pred. No. 4.4e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 240
```

RESULT 6
AAB67805
ID AAB67805 standard; Protein; 412 AA.
XX
AC AAB67805;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of G-protein coupled receptor IGS4B short version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;

KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAF80325.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 18; Page 89-90; 102pp; English.
XX
CC The present sequence represents the short version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral.nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 412 AA;

Query Match 99.1%; Score 1252; DB 22; Length 412;
Best Local Similarity 99.6%; Pred. No. 4.4e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180

Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
```

Db 181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 240

RESULT 7
AAY71296
ID AAY71296 standard; Protein; 415 AA.
XX
AC AAY71296;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human orphan G protein-coupled receptor hRUP6.
XX
KW Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
KW transmembrane receptor; signal cascade.
XX
OS Homo sapiens.
XX
PN WO200031258-A2.
XX
PD 02-JUN-2000.
XX
PF 13-OCT-1999; 99WO-US23687.
XX
PR 20-NOV-1998; 98US-0109213.
PR 16-FEB-1999; 99US-0120416.
PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123949.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0136567.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 29-JUN-1999; 99US-0141448.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156634.
PR 29-SEP-1999; 99US-0156653.
PR 01-OCT-1999; 99US-0157280.
PR 01-OCT-1999; 99US-0157281.
PR 01-OCT-1999; 99US-0157282.
PR 01-OCT-1999; 99US-0157293.
PR 01-OCT-1999; 99US-0157294.
PR 12-OCT-1999; 99US-0416760.
PR 12-OCT-1999; 99US-0417044.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Dang HT, Liaw CW, Lin I;
XX
DR WPI; 2000-400068/34.
DR N-PSDB; AAD01123.
XX
PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
PT for use in the identification of G protein-coupled receptor agonists -
XX
PS Claim 22; Page 57-59; 102pp; English.
XX
CC The present amino acid sequence is the hRUP6, an endogenous human
CC orphan G protein-coupled receptor (GPCR). The full length hRUP6 cDNA
CC was cloned by RT-PCR using human thymus cDNA as template.
CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
CC alpha helices with an extracellular N-terminus and an intracellular
CC C-terminus. However, no endogenous ligands has yet been identified for
CC the proteins of the invention. The orphan GPCRs may be used in the
CC identification of their endogenous ligands, and to screen potential GPCR
CC agonists and antagonists for use as pharmaceutical agents. The proteins
CC may also be used in the study of GPCR-mediated signalling cascades, and
CC to elucidate their precise role in normal and diseased human conditions.
CC Nucleic acid encoding human orphan GPCRs may be used for tissue
CC localisation expression analysis to provide information about their

```
CC   function in healthy and pathological states.
XX
SQ   Sequence   415 AA;

  Query Match             99.1%;  Score 1252;  DB 21;  Length 415;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-141;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 8
AAB02830
ID   AAB02830 standard; Protein; 415 AA.
XX
AC   AAB02830;
XX
DT   22-AUG-2000  (first entry)
XX
DE   Human G protein coupled receptor hRUP6 protein SEQ ID NO:12.
XX
KW   Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW   identification; agonist; screening; therapeutic; pharmaceutical;
KW   mutant.
XX
OS   Homo sapiens.
XX
PN   WO200022131-A2.
XX
PD   20-APR-2000.
XX
PF   13-OCT-1999;   99WO-US24065.
XX
PR   13-OCT-1998;   98US-0170496.
PR   12-NOV-1998;   98US-0108029.
PR   20-NOV-1998;   98US-0109213.
PR   27-NOV-1998;   98US-0110060.
PR   16-FEB-1999;   99US-0120416.
PR   26-FEB-1999;   99US-0121852.
PR   12-MAR-1999;   99US-0123944.
PR   12-MAR-1999;   99US-0123945.
PR   12-MAR-1999;   99US-0123946.
PR   12-MAR-1999;   99US-0123948.
PR   12-MAR-1999;   99US-0123949.
PR   12-MAR-1999;   99US-0123951.
PR   28-MAY-1999;   99US-0136436.
PR   28-MAY-1999;   99US-0136437.
PR   28-MAY-1999;   99US-0136439.
PR   28-MAY-1999;   99US-0137127.
PR   28-MAY-1999;   99US-0137131.
PR   28-MAY-1999;   99US-0137567.
PR   30-JUN-1999;   99US-0141448.
PR   27-AUG-1999;   99US-0151114.
PR   03-SEP-1999;   99US-0152524.
PR   29-SEP-1999;   99US-0156633.
PR   29-SEP-1999;   99US-0156555.
PR   29-SEP-1999;   99US-0156634.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP,  Lehmann-Bruinsma K,  Chalmers DT,  Chen R,  Dang HT;
PI   Gore M,  Liaw CW,  Lin I,  Lowitz K,  White C;
XX
DR   WPI; 2000-317986/27.
DR   N-PSDB; AAA46022.
XX
PT   Non-endogenous, human G protein-coupled receptors for screening
PT   receptor, inverse or partial agonists useful as therapeutic agents -
XX
PS   Example 1; Page 86-88; 187pp; English.
XX
CC   The present invention describes transmembrane receptors, preferably
CC   human G protein coupled receptors (GPCR), for which the endogenous
CC   ligand is unknown (orphan GPCR receptors). More specifically the present
CC   invention relates to non-endogenous, constitutively activated versions
CC   of a human GPCR. These non-endogenous human GPCRs can be useful for
CC   the direct identification of candidate compounds as receptors agonists,
CC   inverse agonists or partial agonists for use as pharmaceutical agents.
CC   AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence   415 AA;

  Query Match             99.1%;  Score 1252;  DB 21;  Length 415;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-141;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 9
AAY52992
ID   AAY52992 standard; Protein; 415 AA.
XX
AC   AAY52992;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Human neurotensin-like receptor.
XX
KW   Human; neurotensin-like receptor; NLR: G-protein coupled receptor;
KW   central nervous system; anasthesia; analgesia.
XX
OS   Homo sapiens.
XX
PN   WO9955732-A1.
XX
PD   04-NOV-1999.
XX
PF   15-APR-1999;   99WO-SE00598.
XX
PR   24-APR-1998;   98SE-0001455.
XX
PA   (ASTR-) ASTRA PHARMA INC.
PA   (ASTR ) ASTRA AB.
XX
PI   Ahmad S,  Cao J,  O'Donnell D,  Walker P;
XX
```

DR WPI; 2000-052803/04.
DR N-PSDB; AAZ33297.
XX
PT Novel neurotensin-like receptor, useful for identifying agents for
PT producing anasthesia or analgesia -
XX
PS Claim 1; Fig 2; 46pp; English.
XX
CC The present sequence represents a human G-protein coupled receptor
CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein
CC can be used to isolate compounds that bind, (ant)agonise or alter
CC the activity or expression of the NLR. The NLR is a G-protein coupled
CC receptor which is expressed in the central nervous system and shares
CC homology with human neurotensin receptor. The receptors can be used in
CC assays to identify agents for producing anasthesia and analgesia.
XX
SQ Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 21; Length 415;
Best Local Similarity 99.6%; Pred. No. 4.5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
Db 4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy 61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
Db 64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy 121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
Db 124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy 181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
Db 184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243

RESULT 10
AAG63353
ID AAG63353 standard; Protein; 415 AA.
XX
AC AAG63353;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease.
XX
OS Homo sapiens.
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-JP00746.
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI; 2001-488917/53.
DR N-PSDB; AAH43072.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
PS Claim 1; Page 77-79; 95pp; Japanese.
XX
CC The present sequence represents a human TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventatives and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;
Best Local Similarity 99.6%; Pred. No. 4.5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
Db 4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy 61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
Db 64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy 121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
Db 124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy 181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
Db 184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243

RESULT 11
AAG63366
ID AAG63366 standard; Protein; 415 AA.
XX
AC AAG63366;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease.
XX
OS Homo sapiens.
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-JP00746.
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI; 2001-488917/53.
DR N-PSDB; AAH43075.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX

```
PS   Disclosure; Page 88; 95pp; Japanese.
XX
CC   The present sequence represents a human TGR-1 protein. The specification
CC   describes a method of screening a compound, which is capable of binding
CC   properties of neuromedin U to TGR-1. The method is useful for screening
CC   preventatives and remedies for hypertension, stress diseases, etc..
CC   TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ   Sequence   415 AA;

  Query Match             99.1%;  Score 1252;  DB 22;  Length 415;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-141;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 12
AAG64297
ID   AAG64297 standard; Protein; 415 AA.
XX
AC   AAG64297;
XX
DT   21-SEP-2001  (first entry)
XX
DE   Human GTP-binding protein-coupled receptor GPRv39.
XX
KW   GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
KW   muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
KW   G-protein.
XX
OS   Homo sapiens.
XX
PN   WO200148189-A1.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-JP09409.
XX
PR   28-DEC-1999;   99JP-0375152.
PR   31-MAR-2000; 2000JP-0101339.
PR   23-MAY-2000; 2000JP-0155978.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Matsumoto S,  Oda T,  Saito Y,  Morikawa N,  Yoshida K,  Suwa M;
PI   Sugiyama T;
XX
DR   WPI; 2001-425663/45.
DR   N-PSDB; AAH49526.
XX
PT   Family of guanosine triphosphate binding protein coupled receptors and
PT   genes encoding them for treatment and prevention of diseases associated
PT   with these receptors  -
XX
PS   Claim 1; Pages 100-103; 137pp; Japanese.
XX
CC   The present sequence is the protein sequence for a human guanosine
CC   triphosphate (GTP)-binding protein-coupled receptor. The receptor is
CC   useful for the investigation, diagnosis, treatment and prevention of
CC   diseases associated with GTP-binding protein-coupled receptors, including
CC   neurological, circulatory, digestive system, immune system, muscle and
CC   urinary system disorders. GTP-binding proteins are also known as
CC   G-proteins.
XX
SQ   Sequence   415 AA;

  Query Match             99.1%;  Score 1252;  DB 22;  Length 415;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-141;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 13
AAE03629
ID   AAE03629 standard; Protein; 415 AA.
XX
AC   AAE03629;
XX
DT   07-AUG-2001  (first entry)
XX
DE   Human G-protein coupled receptor, SNORF72.
XX
KW   Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
KW   NMU; inflammation; arthritis; autoimmune disease; septicaemia; psychotic;
KW   mental retardation; transplant rejection; neurological disorder; anxiety;
KW   respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
KW   gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
KW   ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
KW   dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
KW   Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
KW   cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
KW   Cushing's disease; dysmenorrhoea; antianginal; cytostatic; osteoporosis;
KW   metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
KW   tranquiliser; antiulcer; antiaddictive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          44..71
FT                   /label= Transmembrane_domain
FT   Domain          82..98
FT                   /label= Transmembrane_domain
FT   Domain          124..142
FT                   /label= Transmembrane_domain
FT   Domain          166..179
FT                   /label= Transmembrane_domain
FT   Domain          215..240
FT                   /label= Transmembrane_domain
FT   Domain          269..286
FT                   /label= Transmembrane_domain
FT   Domain          301..330
FT                   /label= Transmembrane_domain
XX
PN   WO200144297-A1.
XX
PD   21-JUN-2001.
```

XX
PF 13-DEC-2000; 2000WO-US33787.
XX
PR 17-DEC-1999; 99US-0466435.
PR 25-APR-2000; 2000US-0558099.
PR 30-JUN-2000; 2000US-0609146.
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
XX
PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
XX
DR WPI; 2001-390240/41.
DR N-PSDB; AAD08008.
XX
PT A purified mammalian SNORF62 or SNORF72 receptor protein for
PT identification of compounds to treat e.g. inflammation, arthritis,
PT autoimmune diseases, transplant rejection, AIDS, cancer -
XX
PS Claim 12; Fig 4; 256pp; English.
XX
CC The invention relates to human G-protein coupled receptors, SNORF62 and
CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
CC are also known as NMU receptors. The agonist and antagonist of NMU
CC receptors are useful for treating an abnormality in a subject that is
CC alleviated by decreasing or increasing the activity of NMU receptor.
CC The NMU receptors serves as a valuable tool for designing drugs which are
CC useful for treating various pathophysiological conditions such as
CC inflammation, arthritis, autoimmune diseases, transplant rejection,
CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC septicaemia, AIDS, pain, psychotic and neurological disorders, including
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC eating/body weight disorders including obesity, bulimia, diabetes,
CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC disorders, ischaemia, stroke, cancers, sexual disorders, circadian
CC disorders, renal disorders, bone diseases including osteoporosis, benign
CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC dermatological disorders such as psoriasis, allergies, Parkinson's
CC disease, Alzheimer's disease, acute heart failure, angina disorders,
CC delirium and dyskinesias such as Huntington's disease. They can also be
CC used to regulate steroid hormone disorders, epinephrine release
CC disorders, electrolyte balance disorders, endocrine disorders, memory
CC disorders, somatosensory disorders, metabolic disorders, behavioural
CC disorders, drug addiction, migraine, Addison's disease, Cushing's
CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
CC The present sequence is human G-protein coupled receptor, SNORF72.
XX
SQ Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;
Best Local Similarity 99.6%; Pred. No. 4.5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243
```

RESULT 14
AAB67802

ID AAB67802 standard; Protein; 415 AA.
XX
AC AAB67802;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of G-protein coupled receptor IGS4A long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAF80322.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 18; Page 77-79; 102pp; English.
XX
CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;
Best Local Similarity 99.6%; Pred. No. 4.5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63
```

```
Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243
```

RESULT 15
AAB67804
ID AAB67804 standard; Protein; 415 AA.
XX
AC AAB67804;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a G-protein coupled receptor IGS4B long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
XX
OS Homo sapiens.
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI; 2001-273568/28.
DR N-PSDB; AAF80324.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 19; Page 85-86; 102pp; English.
XX
CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;
Best Local Similarity 99.6%; Pred. No. 4.5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243
```

Search completed: January 17, 2003, 05:04:03
Job time : 83 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 04:30:56 ; Search time 41 Seconds
(without alignments)
567.428 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	338	26.8	424	2	JH0164	neurotensin recept
2	332.5	26.3	378	2	T15816	hypothetical prote
3	325.5	25.8	418	2	S29506	neurotensin recept
4	303.5	24.0	418	2	A88013	protein K10B4.4 [i
5	281	22.2	416	2	S68822	neurotensin recept
6	281	22.2	564	2	A38271	serotonin receptor
7	265	21.0	393	2	A39251	thyrotropin-releas
8	259	20.5	398	2	JN0708	thyrotropin-releas
9	258	20.4	411	2	I56444	thyrotrophin-relea
10	258	20.4	412	2	S23436	thyroliberin recep
11	257.5	20.4	402	2	I56595	neurokinin 2 recep
12	255.5	20.2	392	2	S65693	opioid receptor mu
13	255	20.2	398	2	A57510	mu opioid receptor
14	253.5	20.1	400	2	I56553	mu opiate receptor
15	252.5	20.0	369	2	B41795	somatostatin recep
16	248	19.6	369	2	JC2083	somatostatin recep
17	247.5	19.6	369	2	A45291	somatostatin recep
18	243.5	19.3	384	2	S20303	neurokinin 2 recep
19	243	19.2	363	2	I57940	somatostatin recep
20	242.5	19.2	385	2	S55524	neurokinin 3 recep
21	242.5	19.2	398	2	I56517	mu opioid receptor
22	241.5	19.1	384	1	S00516	neurokinin 2 recep
23	241	19.1	359	2	S15403	angiotensin II rec
24	240.5	19.0	442	1	DYXLD2	dopamine receptor
25	240	19.0	359	2	JH0621	angiotensin II rec
26	240	19.0	359	2	JC2134	angiotensin II rec
27	239.5	19.0	345	2	T24659	hypothetical prote
28	239	18.9	359	2	JC1104	angiotensin II rec
29	239	18.9	367	2	I49022	kappa opioid recep
30	239	18.9	367	2	JC2421	opioid receptor ho
31	239	18.9	367	2	I56520	G protein-coupled
32	238.5	18.9	346	2	S29248	somatostatin recep
33	238.5	18.9	369	2	D41795	somatostatin recep
34	238.5	18.9	370	2	S43087	orphan opioid rece
35	238.5	18.9	452	2	A34916	neurokinin 3 recep
36	238	18.8	359	2	S44425	angiotensin II rec
37	238	18.8	359	2	A42656	angiotensin II rec
38	237.5	18.8	398	1	JQ1059	neurokinin 2 recep
39	237.5	18.8	511	2	C56849	dopamine receptor-
40	237	18.8	444	1	DYMSD2	dopamine receptor
41	237	18.8	444	1	S08146	dopamine receptor
42	237	18.8	460	2	A32605	serotonin receptor
43	236.5	18.7	387	2	JC5949	galanin receptor 2
44	235	18.6	359	2	A48857	angiotensin II rec
45	234	18.5	359	2	JC1194	angiotensin II rec

ALIGNMENTS

RESULT 1
JH0164
neurotensin receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C;Accession: JH0164
R;Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A;Title: Structure and functional expression of the cloned rat neurotensin receptor.
A;Reference number: JH0164; MUID:90297956; PMID:1694443
A;Accession: JH0164
A;Molecule type: mRNA
A;Residues: 1-424 <TAN>
C;Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
tter (neuromodulator in the brain and as a hormone) cellular mediator in peripheral t
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;65-87/Domain: transmembrane #status predicted <TM1>
F;97-121/Domain: transmembrane #status predicted <TM2>
F;144-165/Domain: transmembrane #status predicted <TM3>
F;189-210/Domain: transmembrane #status predicted <TM4>
F;236-260/Domain: transmembrane #status predicted <TM5>
F;309-330/Domain: transmembrane #status predicted <TM6>
F;348-372/Domain: transmembrane #status predicted <TM7>
F;4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.8%; Score 338; DB 2; Length 424;
Best Local Similarity 35.1%; Pred. No. 1.3e-22;
Matches 72; Conservative 50; Mismatches 71; Indels 12; Gaps 5;

```
Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY 101
           |: :|: :|||| :|| :    : :    |::::  :|:| |||:||||:||| ||:|:|
Db      67 VTAIYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHLGSLALSDLLILLLAMPVELY 126

Qy     102 E-MWRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR 157
             :| ::|: ||  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :|
Db     127 NFIWVHHPWAFGDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSR 183

Qy     158 RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVT 217
           |  : :  :|  | | ::|     |::          ||   || |        :|||
Db     184 SRTKKFISAIWLASALLAIPMLFTMGLQNR--SGDGTHPGGLVCTPIVDTATVKVVIQVN 241

Qy     218 SFLFYLLPMTVISVLYYLMALRVSI 242
           :|: :| || |||:|  ::| ::::
Db     242 TFMSFLFPMLVISILNTVIANKLTV 266
```

RESULT 2
T15816
hypothetical protein C48C5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000

```
C;Accession: T15816
R;Favello, A.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C48C5.
A;Reference number: Z18410
A;Accession: T15816
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-378 <FAV>
A;Cross-references: EMBL:U39994; NID:g1055102; PID:g1055105; PIDN:AAB37017.1; GSPDB:GN00
A;Experimental source: strain Bristol N2; clone C48C5
C;Genetics:
A;Gene: CESP:C48C5.1
A;Map position: X
A;Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
C;Superfamily: adenosine receptor A1

  Query Match             26.3%;  Score 332.5;  DB 2;  Length 378;
  Best Local Similarity   33.6%;  Pred. No. 3.6e-22;
  Matches   87;  Conservative   37;  Mismatches   98;  Indels   37;  Gaps    6;

Qy      20 QKHLNSTEEY---LAFLCGPRRSH----------------------FFLPVSVVYVPIF 53
           |  ||:||:    ||| |    ||                          |: :|: ||
Db       3 QACLNTTEDQCDCLAFNCPIVYSHSESEKEACYMEHCFISKRALDDVTLYKVTALYIFIF 62

Qy      54 VVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGP 113
           :||||||   |||: :|  |||  : || :||||||: | :|:| ||   |  ||: |
Db      63 LVGVIGNTTTCLVMKKHPMMKTHASMYLMNLAVSDLVTLCVGLPFEVMMNWNQYPWPFPD 122

Qy     114 VGCYFKTALFETVCFASILSITTVSVERYVAILHP-FRAKLQSTRRRALRILGIVWGFSV 172
             |  |  : ||    |||:|   ::||||||: || |  |:|  :|    |:|  | ||:
Db     123 YICNLKALIAETTSSVSILTILIFAIERYVAVCHPLFLMKVQPFKRNIGTIIGFTWIFSI 182

Qy     173 LFSLPNTSIHGIKF--HYFP---NGSLVPGSATCTVI-----KPMWIYNFIIQVTSFLFY 222
           | ::|    |   :    :|   |   |  | | :       |    :  :   ::  |:
Db     183 LCAMPFAIHHRADYIMKSWPGTDNRIPVKSSKMCMIAVMFEPKLASTFKILFHFSAIAFF 242

Qy     223 LLPMTVISVLYYLMALRVS 241
             ||:  | :||  :| :||
Db     243 ALPLFTIVILYARIACKVS 261


RESULT 3
S29506
neurotensin receptor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C;Accession: S29506
R;Vita, N.; Laurent, P.; Lefort, S.; Chalon, P.; Dumont, X.; Kaghad, M.; Gully, D.; le F
FEBS Lett. 317, 139-142, 1993
A;Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
A;Reference number: S29506; MUID:93154505; PMID:8381365
A;Accession: S29506
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-418 <VIT>
A;Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             25.8%;  Score 325.5;  DB 2;  Length 418;
  Best Local Similarity   32.6%;  Pred. No. 1.7e-21;
  Matches   86;  Conservative   49;  Mismatches   76;  Indels   53;  Gaps   10;

Qy      17 DPFQKHLNSTEEYL--------------AFLCGPRRS-------HFFLPVSVVYVPIFVV 55
           ||||:     || |               |  |          :  : |: ||: :|||
Db      17 DPFQRAQAGLEEALLAPGFGNASGNASERVLAAPSSELDVNTDIYSKVLVTAVYLALFVV 76

Qy      56 GVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE-MWRNYPFLF 111
           | :|| :    : :    |:::  :|:| |||:|||| ||| ||:|:|  :| ::|: |
Db      77 GTVGNTVTAFTLARKKSLQSLQSTVHYHLGSLALSDLLTLLLAMPVELYNFIWVHHPWAF 136

Qy     112 GPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVW 168
           |  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :| |  : :  :|
Db     137 GDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSRSRTKKFISAIW 193

Qy     169 GFSVLFSLP--------NTSIHGIKFHYFPNGSLV--PGSATCTVIKPMWIYNFIIQVTS 218
             | | ::|        | |  |        | ||  |   | ||         :||| :
Db     194 LASALLTVPMLFTMGEQNRSADGQH-----AGGLVCTPTIHTATV-------KVVIQVNT 241

Qy     219 FLFYLLPMTVISVLYYLMALRVSI 242
           |: :: || |||||  ::| ::::
Db     242 FMSFIFPMVVISVLNTIIANKLTV 265


RESULT 4
A88013
protein K10B4.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: A88013
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: A88013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-418 <STO>
A;Cross-references: GB:chr_II; PIDN:AB71009.1; PID:g2429475; GSPDB:GN00020; CESP:K10B
A;Note: similar to family 1 of G-protein coupled receptors
C;Genetics:
A;Gene: K10B4.4
A;Map position: 2

  Query Match             24.0%;  Score 303.5;  DB 2;  Length 418;
  Best Local Similarity   30.2%;  Pred. No. 1.6e-19;
  Matches   76;  Conservative   54;  Mismatches   89;  Indels   33;  Gaps    6;

Qy      22 HLNSTEEYLAFLCGPR--RSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNY 79
           :::   ||:    | |   :   :|  ::|  ||::|: ||:  |:||  :::|  ||||
Db       9 NVSEITEYVLSTLGERCQSAGIVIPTVIIYGTIFLLGLFGNICTCIVIAANKSMHNPTNY 68

Qy      80 YLFSLAVSDLLVLLLGMPLEVYE-MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVS 138
           |||||||||:: |:||:|:| |: :  :||: |    |  :  | |   :|||: |   |
Db      69 YLFSLAVSDIIALILGLPMEFYQSLDYSYPYRFSEGICKARAFLIEFTSYASIMIICCFS 128

Qy     139 VERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGI-------------- 184
           ||::|| || |:|: ||  ||  :: : |  | : :||   |  |
Db     129 FERWLAICHPLRSKIFSTLWRANVLIILAWTISFVCALPIAFIVQINKLPLPEDAKYQPW 188

Qy     185 --KFHYFP----NGSLVPGSAT---------CTVIKPMWIYNFIIQVTSF-LFYLLPMTV 228
             |  :|      |  : | |           | : :       :| : :| :|:::|
Db     189 TNKVSFFAVGVLNNRIFPVSTDGIFVLHTEFCAMNQSRPDQQKMIIIFAFTVFFVIPAIA 248

Qy     229 ISVLYYLMALRV 240
           | ::|  :|:::
Db     249 IVIMYAHIAVQL 260


RESULT 5
S68822
neurotensin receptor 2, levocabastine-sensitive - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
C;Accession: S68822
R;Chalon, P.; Vita, N.; Kaghad, M.; Guillemot, M.; Bonnin, J.; Delpech, B.; le Fur, G
FEBS Lett. 386, 91-94, 1996
A;Title: Molecular cloning of a levocabastine-sensitive neurotensin binding site.
A;Reference number: S68822; MUID:96228041; PMID:8647296
A;Accession: S68822
A;Molecule type: mRNA
A;Residues: 1-416 <CHA>
A;Cross-references: GB:X97121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580
```

```
A;Experimental source: hypothalamus
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;33-58/Domain: transmembrane #status predicted <TM1>
F;70-91/Domain: transmembrane #status predicted <TM2>
F;110-131/Domain: transmembrane #status predicted <TM3>
F;155-175/Domain: transmembrane #status predicted <TM4>
F;204-230/Domain: transmembrane #status predicted <TM5>
F;296-315/Domain: transmembrane #status predicted <TM6>
F;335-361/Domain: transmembrane #status predicted <TM7>

  Query Match             22.2%;  Score 281;  DB 2;  Length 416;
  Best Local Similarity   35.3%;  Pred. No. 1.6e-17;
  Matches   72;  Conservative   40;  Mismatches   78;  Indels   14;  Gaps    6;

Qy      46 SVVYVPIFVVGVIGNVLVCLVILQHQAMKT-PTNYYLFSLAVSDLLVLLLGMPLEVYE-M 103
           : :|  ||  |  || |   |:|: :| :     |:: |||:| ||:||: ||:|:|  :
Db      36 TALYSLIFAFGTAGNALSVHVVLKARAGRPGRLRYHVLSLALSALLLLLVSMPMELYNFV 95

Qy     104 WRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA 160
           | :||::|| :||   ||   : |   :|::||: ::| || :|:  | ||:   | ||
Db      96 WSHYPWVFGDLGCRGYYF---VRELCAYATVLSVASLSAERCLAVCQPLRARRLLTPRRT 152

Qy     161 LRILGIVWGFSVLFSLPNTSIHGIKFHY-FPNGSLVPGSATCTVIKPMWIYNFIIQVTSF 219
            |:| :||  |:  :||   | | |      :|   | |  |||:          |||
Db     153 RRLLSLVWVASLGLALPMAVIMGQKHEVESADGEPEPASRVCTVLVSRATLQVFIQVNVL 212

Qy     220 LFYLLPMTVISVL-----YYLMAL 238
           : : ||: : : |      :||||
Db     213 VSFALPLALTAFLNGITVNHLMAL 236


RESULT 6
A38271
serotonin receptor 7 - fruit fly (Drosophila melanogaster)
N;Alternate names: 5-hydroxytryptamine receptor 7 (5-HTR7)
C;Species: Drosophila melanogaster
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 04-Sep-1998
C;Accession: A38271
R;Witz, P.; Amlaiky, N.; Plassat, J.L.; Maroteaux, L.; Borrelli, E.; Hen, R.
Proc. Natl. Acad. Sci. U.S.A. 87, 8940-8944, 1990
A;Title: Cloning and characterization of a Drosophila serotonin receptor that activates
A;Reference number: A38271; MUID:91062395; PMID:2174167
A;Accession: A38271
A;Molecule type: mRNA
A;Residues: 1-564 <WIT>
A;Cross-references: GB:M55533
A;Note: the authors translated the codon CTT for residue 213 as Ala, GTG for residue 215
C;Genetics:
A;Gene: FlyBase:5-HT7
A;Cross-references: FlyBase:FBgn0004573
C;Superfamily: octopamine receptor type I
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein; phosphoprotein

  Query Match             22.2%;  Score 281;  DB 2;  Length 564;
  Best Local Similarity   36.6%;  Pred. No. 2.1e-17;
  Matches   79;  Conservative   36;  Mismatches   73;  Indels   28;  Gaps    9;

Qy      27 EEYLA--FLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSL 84
           ||: |  |:  |  | |   ||:| : : :  |:||||||: :   : :: | || | ||
Db     148 EEFAAGEFVLRPLTSIF---VSIVLLIVILGTVVGNVLVCIAVCMVRKLRRPCNYLLVSL 204

Qy      85 AVSDLLVLLLGMPL----EVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVE 140
           |:||| | || ||:    || | |      |||: |    :     | ||||::  :||:
Db     205 ALSDLCVALLVMPMALLYEVLEKWN-----FGPLLCDIWVSFDVLCCTASILNLCAISVD 259

Qy     141 RYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSAT 200
           ||:||  |    :: | || :  :|||| :   |||  | | : |    |  :
Db     260 RYLAITKPLEYGVKRTPRRMMLCVGIVWLAAACISLPPLLILGNE-HEDEEGQPI----- 313

Qy     201 CTVIKPMWIYNFIIQVTSFL--FYLLPMTVISVLYY 234
           ||| :    ||  |: : |  || :|::|:  :||
Db     314 CTVCQ-----NFAYQIYATLGSFY-IPLSVMLFVYY 343


RESULT 7
A39251
thyrotropin-releasing hormone receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 11-Jan-2000
C;Accession: A39251
R;Straub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990
A;Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releas
A;Reference number: A39251; MUID:91088548; PMID:2175902
A;Accession: A39251
A;Molecule type: mRNA
A;Residues: 1-393 <STR>
A;Cross-references: GB:M59811; GB:M37490; NID:g202153; PIDN:AAA40480.1; PID:g202154
C;Superfamily: adenosine receptor A1
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

  Query Match             21.0%;  Score 265;  DB 2;  Length 393;
  Best Local Similarity   28.9%;  Pred. No. 4e-16;
  Matches   72;  Conservative   48;  Mismatches   69;  Indels   60;  Gaps    9;

Qy      15 LEDPFQKHLNSTE---------EY-------LAFLCGPRRSHFFLPVSVVYVPIFVVGVI 58
           :|:     :| ||         ||       :  :||                  :|::
Db       1 MENDTVSEMNQTELQPQAAVALEYQVVTILLVVIICG------------------LGIV 41

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---EVYEMWRNYPFLFGPV 114
           ||::| ||::: : |:|||| || ||||:||:||:  |:|     :|  |     ::| |
Db      42 GNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDSIYGSW-----VYGYV 96

Qy     115 GCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLF 174
           ||   | |      ||  |||  ::|||:|| || :|:   |  || :|:  || |: ::
Db      97 GCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWAFTSIY 156

Qy     175 SLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFIIQVTSFLFYLLPMTV 228
           :    :  :    : |  :|    :|          |::: :| :      ||::|| :
Db     157 CMLWFFLLDLNISTYKNAVVV----SCGYKISRNYYSPIYLMDFGV------FYVVPMIL 206

Qy     229 ISVLYYLMA 237
           :|||  :|
Db     207 ATVLYGFIA 215


RESULT 8
JN0708
thyrotropin-releasing hormone receptor - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C;Accession: S40682; JN0759; S50151; S50152; I38356; JN0708
R;Matre, V.; Karlsen, H.E.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O
Biochem. Biophys. Res. Commun. 195, 179-185, 1993
A;Title: Molecular cloning of a human thyrotropin-releasing hormone recept
A;Reference number: S40682; MUID:93371401; PMID:8395824
A;Accession: S40682
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-398 <MAT>
A;Cross-references: EMBL:X75071; NID:g404157; PIDN:CAA52965.1; PID:g404158
R;Yamada, M.; Monden, T.; Satoh, T.; Satoh, N.; Murakami, M.; Iriuchijima, T.; Kakega
Biochem. Biophys. Res. Commun. 195, 737-745, 1993
A;Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing
ptor gene.
A;Reference number: JN0759; MUID:93384596; PMID:8396925
A;Accession: JN0759
A;Molecule type: mRNA
A;Residues: 1-398 <YAM>
A;Cross-references: GB:D16845; NID:g577631; PIDN:BAA04120.1; PID:g577632
R;Hinuma, S.; Hosoya, M.; Ogi, K.; Tanaka, H.; Nagai, Y.; Onda, H.
Biochim. Biophys. Acta 1219, 251-259, 1994
A;Title: Molecular cloning and functional expression of a human thyrotropin-releasing
A;Reference number: S50151; MUID:95002135; PMID:7918619
A;Accession: S50151
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <HIN>
A;Accession: S50152
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 267-398 <HI2>
R;Duthie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eidne, K.A.
Mol. Cell. Endocrinol. 95, R11-R15, 1993
A;Title: Cloning and functional characterisation of the human TRH receptor.
A;Reference number: I38356; MUID:94063224; PMID:8243797
A;Accession: I38356
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-398 <RES>
A;Cross-references: EMBL:X72089; NID:g440155; PIDN:CAA50979.1; PID:g440156
C;Genetics:
A;Gene: GDB:TRHR
A;Cross-references: GDB:228955; OMIM:188545
A;Map position: 8q23-8q23
C;Superfamily: adenosine receptor A1
C;Keywords: G protein-coupled receptor; receptor; transmembrane protein
F;29-51/Domain: transmembrane #status predicted <TM1>
F;62-83/Domain: transmembrane #status predicted <TM2>
F;101-121/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;194-215/Domain: transmembrane #status predicted <TM5>
F;267-288/Domain: transmembrane #status predicted <TM6>
F;297-319/Domain: transmembrane #status predicted <TM7>

  Query Match             20.5%;  Score 259;  DB 2;  Length 398;
  Best Local Similarity   32.4%;  Pred. No. 1.4e-15;
  Matches   66;  Conservative   47;  Mismatches   65;  Indels   26;  Gaps    7;

Qy      45 VSVVYVPIFV-VGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---E 99
           |::: | |   :|::||::| ||::: : |:|||| || ||||:||:||:  |:|
Db      27 VTILLVLIICGLGIVGNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDS 86

Qy     100 VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRR 159
           :|  |     ::| |||   | |      || |||  ::|||:|| || :|:   | |
Db      87 IYGSW-----VYGYVGCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSR 141

Qy     160 ALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFI 213
           | :|:  || |: |: :   :  :  : :  ::    :|          |::: :|
Db     142 AKKIIIFVWAFTSLYCMLWFFLLDLNISTYKDAIVI----SCGYKISRNYYSPIYLMDFG 197

Qy     214 IQVTSFLFYLLPMTVISVLYYLMA 237
           :      ||::|| : :|||  :|
Db     198 V------FYVVPMILATVLYGFIA 215


RESULT 9
I56444
Thyrotrophin-releasing hormone receptor - mouse
Species: Mus sp. (mouse)
Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000
C;Accession: I56444
R;Sellar, R.E.; Taylor, P.L.; Lamb, R.F.; Zabavnik, J.; Anderson, L.; Eidne, K.A.
J. Mol. Endocrinol. 10, 199-206, 1993
A;Title: Functional expression and molecular characterization of the thyrotrophin-releas
A;Reference number: I56444; MUID:93249585; PMID:8387312
A;Accession: I56444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-411 <RES>
A;Cross-references: GB:S60053; NID:g300151; PIDN:AAB26491.1; PID:g300152
C;Superfamily: adenosine receptor A1

  Query Match             20.4%;  Score 258;  DB 2;  Length 411;
  Best Local Similarity   30.4%;  Pred. No. 1.8e-15;
  Matches   72;  Conservative   49;  Mismatches   80;  Indels   36;  Gaps    8;

Qy      15 LEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVV----GVIGNVLVCLVILQH 70
           :|:     || ||        | :    |   || : : ||    |::||::| ||:::
Db       1 MENETVSELNQTEL-------PPQVAVALEYQVVTILLVVVICGLGIVGNIMVVLVVMRT 53

Qy      71 QAMKTPTNYYLFSLAVSDLLVLL-LGMPL---EVYEMWRNYPFLFGPVGCYFKTALFETV 126
           : |:| || || ||||:||:||:  |:|     :|  |     ::| |||   | |
Db      54 KHMRTATNCYLVSLAVADLMVLVAAGLPNITDSIYGSW-----VYGYVGCLCITYLQYLG 108

Qy     127 CFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKF 186
           ||  |||  ::|||:|| || :|:   |  || :|:  || |: :: :   :  :
Db     109 INASSCSITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWAFTSIYCMLWFFLLDLNI 168

Qy     187 HYFPNGSLVPGSATC------TVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA 237
           : :  ::    :|          |::: :| :      ||::|| : :|||  :|
Db     169 STYKDAIVI----SCGYKISRNYYSPIYLMDFGV------FYVMPMILATVLYGFIA 215


RESULT 10
S23436
thyroliberin receptor - rat
N;Alternate names: thyrotropin-releasing hormone receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000
C;Accession: S23436; I53279; A49168; PQ0326
R;de la Pena, P.; Delgado, L.M.; del Camino, D.; Barros, F.
Biochem. J. 284, 891-899, 1992
A;Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH
A;Reference number: S23436; MUID:92322017; PMID:1377915
A;Accession: S23436
A;Molecule type: mRNA
A;Residues: 1-412 <PEN>
A;Cross-references: EMBL:X64630; NID:g57394; PIDN:CAA45913.1; PID:g57395
R;Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
Endocrinology 134, 432-440, 1994
A;Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotrop
A;Reference number: I53279; MUID:94102223; PMID:8275956
A;Accession: I53279
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-412 <RES>
A;Cross-references: GB:D17469; NID:g464199; PIDN:BAA04289.1; PID:g464200
R;Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chin, W.W.; T
Endocrinology 130, 3529-3536, 1992
A;Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thyr
A;Reference number: A49168; MUID:92283212; PMID:1317787
A;Accession: A49168
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12,'D',14-290,'K',292-412 <ZHA>
A;Experimental source: GH cells
A;Note: sequence extracted from NCBI backbone (NCBIN:104788, NCBIP:104795)
R;Yamada, M.; Monden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori,
Biochem. Biophys. Res. Commun. 184, 367-372, 1992
A;Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA level
A;Reference number: PQ0326; MUID:92231953; PMID:1373613
A;Accession: PQ0326
A;Molecule type: mRNA
A;Residues: 30-58,'P',60-222,'T',224-261 <YAM>
A;Experimental source: strain Wister
A;Note: the authors translated the codon ACA for residue 88 as Ala
C;Superfamily: adenosine receptor A1
C;Keywords: G protein-coupled receptor; transmembrane protein
F;1-22/Domain: transmembrane #status predicted <TM1>
F;32-54/Domain: transmembrane #status predicted <TM2>
F;71-92/Domain: transmembrane #status predicted <TM3>
F;116-140/Domain: transmembrane #status predicted <TM4>
F;165-186/Domain: transmembrane #status predicted <TM5>

  Query Match             20.4%;  Score 258;  DB 2;  Length 412;
  Best Local Similarity   30.4%;  Pred. No. 1.8e-15;
  Matches   72;  Conservative   49;  Mismatches   80;  Indels   36;  Gaps    8;

Qy      15 LEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVV----GVIGNVLVCLVILQH 70
           :|:     || ||        | :    |   || : : ||    |::||::| ||:::
```

```
Db      1 MENETVSELNQTEL-------PPQVAVALEYQVVTILLVVVICGLGIVGNIMVVLVVMRT 53

Qy     71 QAMKTPTNYYLFSLAVSDLLVLL-LGMPL---EVYEMWRNYPFLFGPVGCYFKTALFETV 126
          : |:| || || |||||:||:||:  |:|     :|  |     ::| |||   | |
Db     54 KHMRTATNCYLVSLAVADLMVLVAAGLPNITDSIYGSW-----VYGYVGCLCITYLQYLG 108

Qy    127 CFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKF 186
            ||  |||  ::|||:|| || :|:   |  || :|:  || |: :: :    :  :
Db    109 INASSCSITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWAFTSIYCMLWFFLLDLNI 168

Qy    187 HYFPNGSLVPGSATC------TVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA 237
            : :  ::    :|          |::: :| :      ||::|| : :|||  :|
Db    169 STYKDAIVI----SCGYKISRNYYSPIYLMDFGV------FYVMPMILATVLYGFIA 215


RESULT 11
I56595
neurokinin 2 receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
C;Accession: I56595
R;Aharony, D.; Little, J.; Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
J. Recept. Res. 14, 399-421, 1994
A;Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-pig l
A;Reference number: I56595; MUID:95182423; PMID:7877137
A;Accession: I56595
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-402 <RES>
A;Cross-references: GB:S76253; NID:g913274; PIDN:AAB33553.1; PID:g913275
C;Superfamily: neurokinin 1 receptor

  Query Match             20.4%;  Score 257.5;  DB 2;  Length 402;
  Best Local Similarity   28.9%;  Pred. No. 1.9e-15;
  Matches   59;  Conservative   45;  Mismatches   71;  Indels   29;  Gaps    6;

Qy     46 SVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEV----Y 101
          :  |: : :| | ||  |  :|| || |:| |||:: :||::|| :             :
Db     37 ATAYLALVLVAVTGNATVTWIILAHQRMRTVTNYFIVNLALADLCMAAFNAAFNFVYASH 96

Qy    102 EMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRAL 161
           :|      ||   |||:     |  | || |:| ::::||:||:|||: :| :   :|
Db     97 NIW-----YFGRAFCYFQNLFPITAMFVSIYSMTAIAIDRYMAIVHPFQPRLSAPSTKA- 150

Qy    162 RILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVI-------KPMWIYNFII 214
           ::| :|  ::  : |          ::   :   |:  | |        | : :|: ::
Db    151 -VIGGIWLVALALAFPQC--------FYSTITEDEGATKCVVAWPEDSRDKSLLLYHLVV 201

Qy    215 QVTSFLFYLLPMTVISVLYYLMAL 238
           |   | ||||:||: | | :: |
Db    202 IV---LIYLLPLTVMFVAYSIIGL 222


RESULT 12
S65693
opioid receptor mu variant MOR1A - human
C;Species: Homo sapiens (man)
C;Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
C;Accession: S65693; S51216
R;Bare, L.A.; Mansson, E.; Yang, D.
submitted to the EMBL Data Library, July 1994
A;Description: Expression of two variants of the human mu opioid receptor mRNA in SK-N-S
A;Reference number: S65693
A;Accession: S65693
A;Molecule type: mRNA
A;Residues: 1-392 <BAR>
A;Cross-references: EMBL:U12569; NID:g607911; PIDN:AAB60354.1; PID:g607912
R;Bare, L.A.; Mansson, E.; Yang, D.
FEBS Lett. 354, 213-216, 1994
A;Title: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cell
A;Reference number: S51215; MUID:95046336; PMID:7957926
A;Accession: S51216
A;Molecule type: mRNA
A;Residues: 387-392 <BAW>
C;Superfamily: vertebrate rhodopsin

  Query Match             20.2%;  Score 255.5;  DB 2;  Length 392;
  Best Local Similarity   28.7%;  Pred. No. 2.8e-15;
  Matches   73;  Conservative   47;  Mismatches   91;  Indels   43;  Gaps    8;

Qy      8 SWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF-----------------FLPVSVVYV 50
          ||:           :|:  :  |:  |||  |:                     : :  :|
Db     30 SWV----------NLSHLDGDLSDPCGPNRTDLGGRDSLCPPTGSPSMITAITIMALYS 78

Qy     51 PIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPF 109
           : |||: || ||  ||:::  ||| || |:|:||::| |     :| : |  :   :|
Db     79 IVCVVGLFGNFLVMYVIVRYTKMKTATNIYIFNLALADALATST-LPFQSVNYLMGTWP- 136

Qy    110 LFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWG 169
           || : |    ::     | || :: |:|||:||:|: || :|    | | |  |    |
Db    137 -FGTILCKIVISIDYYNMFTSIFTLCTMSVDRYIAVCHPVKALDFRTPRNAKIINVCNWI 195

Qy    170 FSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPM 226
           |    ||   |   |:          ||  ||:    | | :  ::::  |:| :::|:
Db    196 LSSAIGLPVMFIATTKYR--------QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPV 247

Qy    227 TVISVLYYLMALRV 240
           :|:| | || ||:
Db    248 LIITVCYGLMILRL 261


RESULT 13
A57510
mu opioid receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
C;Accession: A57510; I48665; S66513; I49300
R;Kaufman, D.L.; Keith Jr., D.E.; Anton, B.; Tian, J.; Magendzo, K.; Newman, D.; Tran
J. Biol. Chem. 270, 15877-15883, 1995
A;Title: Characterization of the murine mu opioid receptor gene.
A;Reference number: A57510; MUID:95318184; PMID:7797593
A;Accession: A57510
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-398 <KAU>
A;Cross-references: GB:U19380
R;Min, B.H.; Augustin, L.B.; Felsheim, R.F.; Fuchs, J.A.; Loh, H.H.
Proc. Natl. Acad. Sci. U.S.A. 91, 9081-9085, 1994
A;Title: Genomic structure analysis of promoter sequence of a mouse mu opioid recepto
A;Reference number: I48665; MUID:94377496; PMID:8090773
A;Accession: I48665
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-398 <RES>
A;Cross-references: EMBL:U10561; NID:g555696; PIDN:AAB60673.1; PID:g565069
R;Rossi, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.
FEBS Lett. 369, 192-196, 1995
A;Title: Antisense mapping the MOR-1 opioid receptor: evidence for alternative splici
A;Reference number: I49300; MUID:95377399; PMID:7649256
A;Accession: S66513
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-398 <ROS>
A;Cross-references: EMBL:U26915; NID:g1055230; PIDN:AAA81170.1; PID:g1055231
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C;Genetics:
A;Gene: MOR-1
A;Introns: 95/2; 213/1; 386/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmemb

  Query Match             20.2%;  Score 255;  DB 2;  Length 398;
  Best Local Similarity   30.7%;  Pred. No. 3.1e-15;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy     34 CGPRR-----SHFFLP-----------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
```

```
          ||| |     ||   |            :  :|  : |||: || ||  ||:::  |||
Db     43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy     77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
          || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db    103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy    136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
          |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db    160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy    196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
           ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db    212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 14
I56553
mu opiate receptor - human
N;Alternate names: MOR1 protein; opioid receptor mu
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C;Accession: I56553; A38991; S41075; S51215
R;Mestek, A.; Hurley, J.H.; Bye, L.S.; Campbell, A.D.; Chen, Y.; Tian, M.; Liu, J.; Schu
J. Neurosci. 15, 2396-2406, 1995
A;Title: The human mu opioid receptor: modulation of functional desensitization by calci
A;Reference number: I56553; MUID:95198115; PMID:7891175
A;Accession: I56553
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-400 <RES>
A;Cross-references: GB:L29301; NID:g459831; PIDN:AAA73958.1; PID:g459832
R;Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
submitted to GenBank, August 1994
A;Reference number: A38991
A;Accession: A38991
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-50,'N',52-233,'V',235-400 <WAN>
A;Cross-references: GB:L25119; NID:g452072; PIDN:AAA20580.1; PID:g452073
R;Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
FEBS Lett. 338, 217-222, 1994
A;Title: Human mu opiate receptor. cDNA and genomic clones, pharmacologic characterizati
A;Reference number: S41075; MUID:94139928; PMID:7905839
A;Accession: S41075
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-50,'N',52-400 <WA2>
R;Bare, L.A.; Mansson, E.; Yang, D.
FEBS Lett. 354, 213-216, 1994
A;Title: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cell
A;Reference number: S51215; MUID:95046336; PMID:7957926
A;Accession: S51215
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 387-400 <BAR>
C;Genetics:
A;Gene: GDB:OPRM1
A;Cross-references: GDB:137216; OMIM:600018
A;Map position: 6q24-6q25
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;73-96/Domain: transmembrane #status predicted <TMM1>
F;107-132/Domain: transmembrane #status predicted <TMM2>
F;144-165/Domain: transmembrane #status predicted <TMM3>
F;188-208/Domain: transmembrane #status predicted <TMM4>
F;236-257/Domain: transmembrane #status predicted <TMM5>
F;283-304/Domain: transmembrane #status predicted <TMM6>
F;323-342/Domain: transmembrane #status predicted <TMM7>
F;9,12,33,40,48/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match            20.1%;  Score 253.5;  DB 2;  Length 400;
  Best Local Similarity  28.3%;  Pred. No. 4.3e-15;
  Matches   72;  Conservative   48;  Mismatches   91;  Indels   43;  Gaps    8;

Qy       8 SWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF----------------FLPVSVVYV 50
           ||:           :|:  :  |:  ||| |:                     : :  :|
Db      30 SWV----------NLSHLDGNLSDPCGPNRTDLGGRDSLCPPTGSPSMITAITIMALYS 78

Qy      51 PIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPF 109
           : |||: || ||  ||::  ||| || |:|:||::| |     :| : |  :   :|
Db      79 IVCVVGLFGNFLVMYVIVRYTKMKTATNIYIFNLALADALATST-LPFQSVNYLMGTWP- 136

Qy     110 LFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWG 169
           || : |    ::     | || :: |:||:||:|: || :|    | | |  |    |
Db     137 -FGTILCKIVISIDYYNMFTSIFTLCTMSVDRYIAVCHPVKALDFRTPRNAKIINVCNWI 195

Qy     170 FSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPM 226
           |    ||   :  |:          ||  ||:    | | :  ::::  |:| :::|:
Db     196 LSSAIGLPVMFMATTKYR--------QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPV 247

Qy     227 TVISVLYYLMALRV 240
           :|:| | || ||:
Db     248 LIITVCYGLMILRL 261


RESULT 15
B41795
somatostatin receptor 2 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: B41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somat
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: B41795
A;Molecule type: DNA
A;Residues: 1-369 <YAM>
A;Cross-references: GB:M81830; NID:g307435; PIDN:AAA58248.1; PID:g307436
A;Note: sequence extracted from NCBI backbone (NCBIN:74769, NCBIP:74770)
C;Genetics:
A;Gene: GDB:SSTR2
A;Cross-references: GDB:134186; OMIM:182452
A;Map position: 17q24-17q24
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein;
F;44-69/Domain: transmembrane #status predicted <TM1>
F;80-105/Domain: transmembrane #status predicted <TM2>
F;117-138/Domain: transmembrane #status predicted <TM3>
F;158-180/Domain: transmembrane #status predicted <TM4>
F;205-235/Domain: transmembrane #status predicted <TM5>
F;254-281/Domain: transmembrane #status predicted <TM6>
F;288-315/Domain: transmembrane #status predicted <TM7>
F;9,22,29,32,351/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;115-193/Disulfide bonds: #status predicted
F;250/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pre
F;328/Binding site: palmitate (Cys) (covalent) #status predicted

  Query Match            20.0%;  Score 252.5;  DB 2;  Length 369;
  Best Local Similarity  30.6%;  Pred. No. 4.9e-15;
  Matches   68;  Conservative   44;  Mismatches   87;  Indels   23;  Gaps    7;

Qy      24 NSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFS 83
           | || |               :: :|  : ::|: || ||  |||::  ||| || |: :
Db      32 NQTEPYYDLTSN-------AVLTFIYFVVCIIGLCGNTLVIYVILRYAKMKTITNIYILN 84

Qy      84 LAVSDLLVLLLGMP-LEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERY 142
           ||::| |  :||:| | :     ::|  ||   |     :     | ||  :| :|::||
Db      85 LAIADEL-FMLGLPFLAMQVALVHWP--FGKAICRVVMTVDGINQFTSIFCLTVMSIDRY 141

Qy     143 VAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCT 202
           :|::|| ::      | |  |   ||| |:|  ||     |:: : :       | ::||
Db     142 LAVVHPIKSAKWRRPRTAKMITMAVWGVSLLVILPIMIYAGLRSNQW-------GRSSCT 194
```

```
Qy     203 VIKP----MWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           :  |     |   |||  |  | :|:|:|:| : |  : ::|
Db     195 INWPGESGAWYTGFII-YTFILGFLVPLTIICLCYLFIIIKV 235
```

Search completed: January 17, 2003, 05:07:32
Job time : 47 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:39:56 ; Search time 23 Seconds
(without alignments)
436.403 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	338	26.8	424	1	NTR1_RAT	P20789 rattus norv
2	333	26.4	412	1	MTLR_HUMAN	O43193 homo sapien
3	332.5	26.3	424	1	NTR1_MOUSE	O88319 mus musculu
4	325.5	25.8	418	1	NTR1_HUMAN	P30989 homo sapien
5	281	22.2	416	1	NTR2_RAT	Q63384 rattus norv
6	281	22.2	564	1	5HT1_DROME	P20905 drosophila
7	277	21.9	366	1	GHSR_HUMAN	Q92847 homo sapien
8	274	21.7	364	1	GHSR_RAT	O08725 rattus norv
9	274	21.7	366	1	GHSR_PIG	Q95254 sus scrofa
10	270.5	21.4	395	1	TRFR_CHICK	O93603 gallus gall
11	268	21.2	257	1	GHSR_MOUSE	Q99p50 mus musculu
12	265	21.0	393	1	TRFR_MOUSE	P21761 mus musculu
13	259	20.5	398	1	TRFR_HUMAN	P34981 homo sapien
14	258	20.4	398	1	TRFR_BOVIN	O46639 bos taurus
15	258	20.4	398	1	TRFR_SHEEP	Q28596 ovis aries
16	258	20.4	410	1	NTR2_HUMAN	O95665 homo sapien
17	258	20.4	412	1	TRFR_RAT	Q01717 rattus norv
18	257.5	20.4	400	1	OPRM_MACMU	Q9myw9 macaca mula
19	257.5	20.4	402	1	NK2R_CAVPO	Q64077 cavia porce
20	256.5	20.3	401	1	OPRM_BOVIN	P79350 bos taurus
21	255	20.2	398	1	OPRM_MOUSE	P42866 mus musculu
22	253.5	20.1	400	1	OPRM_HUMAN	P35372 homo sapien
23	253	20.0	417	1	NTR2_MOUSE	P70310 mus musculu
24	252.5	20.0	369	1	SSR2_HUMAN	P30874 homo sapien
25	251.5	19.9	401	1	OPRM_PIG	Q95247 sus scrofa
26	250.5	19.8	460	1	OX2R_RAT	P56719 rattus norv
27	248	19.6	369	1	SSR2_PIG	P34994 sus scrofa
28	247.5	19.6	368	1	SSR2_BOVIN	P34993 bos taurus
29	247.5	19.6	369	1	SSR2_RAT	P30680 rattus norv
30	245.5	19.4	370	1	OPRX_CAVPO	P47748 cavia porce
31	243.5	19.3	384	1	NK2R_MOUSE	P30549 mus musculu
32	243	19.2	363	1	SSR5_RAT	P30938 rattus norv
33	242.5	19.2	385	1	NK3R_MOUSE	P47937 mus musculu
34	242.5	19.2	398	1	OPRM_RAT	P33535 rattus norv
35	241.5	19.1	384	1	NK2R_BOVIN	P05363 bos taurus
36	241.5	19.1	444	1	OX2R_CANFA	Q9tup7 canis famil
37	241	19.1	359	1	AG2R_BOVIN	P25104 bos taurus
38	240.5	19.0	442	1	D2D1_XENLA	P24628 xenopus lae
39	240.5	19.0	444	1	OX2R_HUMAN	O43614 homo sapien
40	240	19.0	359	1	AG2R_CAVPO	Q9wv26 cavia porce
41	240	19.0	359	1	AG2R_MOUSE	P29754 mus musculu
42	240	19.0	359	1	AG2R_PIG	P30555 sus scrofa
43	240	19.0	359	1	AG2R_RAT	P25095 rattus norv
44	240	19.0	359	1	AG2S_MOUSE	P29755 mus musculu
45	239.5	19.0	467	1	NK3R_RABIT	O97512 oryctolagus

ALIGNMENTS

RESULT 1
NTR1_RAT
ID NTR1_RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity levocabastine-
DE insensitive neurotensin receptor) (NTRH).
GN NTSR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin
RT receptor.";
RL Neuron 4:847-854(1990).
CC -!- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
DR PIR; JH0164; JH0164.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR003985; NT1_rec.
DR InterPro; IPR003984; NT_rec.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PRINTS; PR01479; NEUROTENSINR.
DR PRINTS; PR01480; NEUROTENSN1R.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.

FT	DOMAIN	1	64 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	65	87 1 (POTENTIAL).
FT	DOMAIN	88	96 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	97	121 2 (POTENTIAL).
FT	DOMAIN	122	143 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	144	165 3 (POTENTIAL).
FT	DOMAIN	166	188 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	189	210 4 (POTENTIAL).
FT	DOMAIN	211	235 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	236	260 5 (POTENTIAL).
FT	DOMAIN	261	308 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	309	330 6 (POTENTIAL).
FT	DOMAIN	331	348 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	349	372 7 (POTENTIAL).
FT	DOMAIN	373	424 CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	4	4 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    142    225       BY SIMILARITY.
FT   LIPID       388    388       PALMITATE (POTENTIAL).
SQ   SEQUENCE   424 AA;  47054 MW;  A9C2F7EAF8D9BCD3 CRC64;

  Query Match             26.8%;  Score 338;  DB 1;  Length 424;
  Best Local Similarity   35.1%;  Pred. No. 1.2e-17;
  Matches   72;  Conservative   50;  Mismatches   71;  Indels   12;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY 101
           |: :|: :|||| :|| :    : :    |:::: :|:| |||:||||:||| ||:|:|
Db      67 VTAIYLALFVVGTVGNSVTAFTLARKKSLQSLQSTVHYHLGSLALSDLLILLLAMPVELY 126

Qy     102 E-MWRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR 157
             :| ::|: ||  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :|
Db     127 NFIWVHHPWAFGDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSR 183

Qy     158 RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVT 217
           |  : :  :|  | | ::|     |::          ||   || |        :|||
Db     184 SRTKKFISAIWLASALLAIPMLFTMGLQNR--SGDGTHPGGLVCTPIVDTATVKVVIQVN 241

Qy     218 SFLFYLLPMTVISVLYYLMALRVSI 242
           :|: :| || |||:|  ::| ::::
Db     242 TFMSFLFPMLVISILNTVIANKLTV 266


RESULT 2
MTLR_HUMAN
ID   MTLR_HUMAN     STANDARD;      PRT;   412 AA.
AC   O43193;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Motilin receptor (G protein-coupled receptor GPR38).
GN   MTLR1 OR MTLR OR GPR38.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=98110578; PubMed=9441746;
RA   McKee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,
RA   Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT   "Cloning and characterization of two human G protein-coupled receptor
RT   genes (GPR38 and GPR39) related to the growth hormone secretagogue
RT   and neurotensin receptors.";
RL   Genomics 46:426-434(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=99316084; PubMed=10381885;
RA   Feighner S.D., Tan C.P., McKee K.K., Palyha O.C., Hreniuk D.L.,
     Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,
     Nargund R., Bakshi R., Abramovitz M., Stocco R., Kargman S.,
     O'Neill G., van Der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,
RA   Howard A.D.;
RT   "Receptor for motilin identified in the human gastrointestinal
RT   system.";
RL   Science 284:2184-2188(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RA   Wall M.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION.
RX   MEDLINE=21219832; PubMed=11322507;
RA   Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
RA   Tan C.P., Mckee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT   "Growth hormone secretagogue receptor family members and ligands.";
RL   Endocrine 14:9-14(2001).
CC   -!- FUNCTION: Receptor for motilin.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE
CC       MARROW.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF034632; AAC26081.1; -.
DR   EMBL; AL137000; CAC19107.1; -.
DR   Genew; HGNC:4495; GPR38.
DR   MIM; 602885; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN       1     35       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    36     56       1 (POTENTIAL).
FT   DOMAIN      57     74       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    75     94       2 (POTENTIAL).
FT   DOMAIN      95    112       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   113    134       3 (POTENTIAL).
FT   DOMAIN     135    157       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   158    178       4 (POTENTIAL).
FT   DOMAIN     179    246       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   247    270       5 (POTENTIAL).
FT   DOMAIN     271    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   299    320       6 (POTENTIAL).
FT   DOMAIN     321    334       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   335    358       7 (POTENTIAL).
FT   DOMAIN     359    412       CYTOPLASMIC (POTENTIAL).
FT   DISULFID   111    235       BY SIMILARITY.
FT   CARBOHYD     6      6       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   301    412       LVVVLAFIICWLPFHVGRIIYINTEDSRMMYFSQYFNIVAL
FT                               QLFYLSASINPILYNLISKKYRAAAFKLLLARKSRPRGFHR
FT                               SRDTAGEVAGDTGGDTVGYTETSANVKTMG -> RKWSRRG
FT                               SKDACLQSAPPGTAQTLGPLPLLAQLWAPLPAPFPISIPAS
FT                               TRRGGGSGIYNLLVALPRWQNHLHKHGRFADDVLLSVL
FT                               (IN ISOFORM B).
SQ   SEQUENCE   412 AA;  45344 MW;  C13FF6165012DEF3 CRC64;

  Query Match             26.4%;  Score 333;  DB 1;  Length 412;
  Best Local Similarity   31.8%;  Pred. No. 2.7e-17;
  Matches   78;  Conservative   51;  Mismatches   70;  Indels   46;  Gaps    5;

Qy      34 CGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL 93
           | |      :||: | : :|||| |||:  ::| ::: |:| || || |:||||||: |
Db      30 CSPFPLGALVPVTAVCLCLFVVGVSGNVVTVMLIGRYRDMRTTTNLYLGSMAVSDLLI-L 88

Qy      94 LGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKL 153
           ||:| ::| :||: |::|||: |     : |   :|::| :| :|||||:||  | ||::
Db      89 LGLPFDLYRLWRSRPWVFGPLLCRLSLYVGEGCTYATLLHMTALSVERYLAICRPLRARV 148

Qy     154 QSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPG--------SATCTVIK 205
             ||||   :: ::|  ::| : |    : |::    |  |:||         |:
Db     149 LVTRRRVRALIAVLWAVALLSAGPFLFLVGVE--QDPGISVVPGLNGTARIASSPLASSP 206

Qy     206 PMWI----------------------------------YNFIIQVTSFLFYLLPMTVISV 231
           |:|:                                    :: ||:  |: ||   :|:
Db     207 PLWLSRAPPPSPPSGPETAEAAALFSRECRPSPAQLGALRVMLWVTTAYFF-LPFLCLSI 265

Qy     232 LYYLM 236
           || |:
Db     266 LTVLM 270
```

```
Db    266 LYGLI 270


RESULT 3
NTR1_MOUSE
ID   NTR1_MOUSE     STANDARD;      PRT;   424 AA.
AC   O88319;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurotensin receptor type 1 (NT-R-1).
GN   NTSR1 OR NTSR.
OS   Mus musculus (Mouse).
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
     NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Snider J., Sano H., Ohta M.;
RT   "Neurotensin receptor type 1.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC       ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC       CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       HIGHEST TO TACHYKININS RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB017027; BAA33013.1; -.
DR   MGD; MGI:97386; Ntsr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003985; NT1_rec.
DR   InterPro; IPR003984; NT_rec.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01479; NEUROTENSINR.
DR   PRINTS; PR01480; NEUROTENSN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     64       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     65     87       1 (POTENTIAL).
FT   DOMAIN       88     96       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     97    121       2 (POTENTIAL).
FT   DOMAIN      122    143       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    144    165       3 (POTENTIAL).
FT   DOMAIN      166    188       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    189    210       4 (POTENTIAL).
FT   DOMAIN      211    235       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    236    260       5 (POTENTIAL).
FT   DOMAIN      261    308       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    309    330       6 (POTENTIAL).
FT   DOMAIN      331    348       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    349    372       7 (POTENTIAL).
FT   DOMAIN      373    424       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    141    224       BY SIMILARITY.
FT   LIPID       388    388       PALMITATE (POTENTIAL).
SQ   SEQUENCE   424 AA;  47216 MW;  8E9A723171A48711 CRC64;

  Query Match           26.3%;  Score 332.5;  DB 1;  Length 424;
  Best Local Similarity  32.1%;  Pred. No. 3e-17;
  Matches   80;  Conservative   55;  Mismatches   85;  Indels   29;  Gaps    7;

Qy      4 LQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLV 63
          | | |    : : :|      :|: |   |           |: ||: :|||| :|| :
Db     36 LSNGSGNSSESILEPNSNLDVNTDIYSKVL----------VTAVYLALFVVGTVGNSVT 84

Qy     64 CLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE-MWRNYPFLFGPVGC--- 116
            : :    |:::: :|:| |||:||||:||| ||:|:|  :| ::|: ||  ||
Db     85 AFTLARKKSLQSLQSTVHYHLGSLALSDLLILLLAMPVELYNFIWVHHPWAFGDAGCRGY 144

Qy    117 YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSL 176
          ||   | :   :|: |:: ::|||||:|| |||:||   :| |  : :  :|  | | ::
Db    145 YF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSRSRTKKFISAIWLASALLAV 201

Qy    177 PNTSIHGIKFHYFPNGSL---VPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLY 233
          |     |::     | |     ||   ||          :||| :|: :| || :||:|
Db    202 PMLFTMGLQ-----NRSADGQHPGGLVCTPTVDTATVKVVIQVNTFMSFLFPMLIISILN 256

Qy    234 YLMALRVSI 242
            ::| ::::
Db    257 TVIANKLTV 265


RESULT 4
NTR1_HUMAN
ID   NTR1_HUMAN     STANDARD;      PRT;   418 AA.
AC   P30989; Q9H4T5; Q9H4H1;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Neurotensin receptor type 1 (NT-R-1) (High-affinity levocabastine-
DE   insensitive neurotensin receptor) (NTRH).
GN   NTSR1 OR NTRR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93154505; PubMed=8381365;
RA   Vita N., Laurent P., Lefort S., Chalon P., Dumont X., Kaghad M.,
RA   Gully D., le Fur G., Ferrara P., Caput D.;
RT   "Cloning and expression of a complementary DNA encoding a high
RT   affinity human neurotensin receptor.";
RL   FEBS Lett. 317:139-142(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
```

```
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC       ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC       CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       HIGHEST TO TACHYKININS RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70070; CAA49675.1; -.
DR   EMBL; AL357033; CAC14923.1; -.
DR   EMBL; AL035669; CAC12747.1; -.
DR   PIR; S29506; S29506.
DR   HSSP; P02699; 1F88.
DR   Genew; HGNC:8039; NTSR1.
DR   MIM; 162651; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003985; NT1_rec.
DR   InterPro; IPR003984; NT_rec.
DR   Pfam; PF00001; 7tm_1; 2.
DR   PRINTS; PR01479; NEUROTENSINR.
DR   PRINTS; PR01480; NEUROTENSN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN       1     63       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     64    86       1 (POTENTIAL).
FT   DOMAIN       87    95       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     96    120      2 (POTENTIAL).
FT   DOMAIN       121   142      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     143   164      3 (POTENTIAL).
FT   DOMAIN       165   187      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     188   209      4 (POTENTIAL).
FT   DOMAIN       210   234      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     235   259      5 (POTENTIAL).
FT   DOMAIN       260   303      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     304   325      6 (POTENTIAL).
FT   DOMAIN       326   343      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     344   367      7 (POTENTIAL).
FT   DOMAIN       368   418      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     4     4        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     37    37       N-LINKED (GLCNAC. . .) (POTENTIAL).
     CARBOHYD     41    41       N-LINKED (GLCNAC. . .) (POTENTIAL).
     DISULFID     141   224      BY SIMILARITY.
     LIPID        383   383      PALMITATE (POTENTIAL).
FT   CONFLICT     200   200      T -> A (IN REF. 2).
SQ   SEQUENCE   418 AA;  46288 MW;  BBBD1EEC2BE6E390 CRC64;

  Query Match           25.8%;  Score 325.5;  DB 1;  Length 418;
  Best Local Similarity  32.6%;  Pred. No. 9.3e-17;
  Matches   86;  Conservative   49;  Mismatches   76;  Indels   53;  Gaps   10;

Qy      17 DPFQKHLNSTEEYL--------------AFLCGPRRS-------HFFLPVSVVYVPIFVV 55
           ||||:     || |              |  |          :  : |: ||: :|||
Db      17 DPFQRAQAGLEEALLAPGFGNASGNASERVLAAPSSELDVNTDIYSKVLVTAVYLALFVV 76

Qy      56 GVIGNVLVCLVILQH---QAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE-MWRNYPFLF 111
           | :|| :    : :    |:::  :|:| |||:|||| ||| ||:|:|  :| ::|: |
Db      77 GTVGNTVTAFTLARKKSLQSLQSTVHYHLGSLALSDLLTLLLAMPVELYNFIWVHHPWAF 136

Qy     112 GPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVW 168
           |  ||   ||   | :   :|: |:: ::|||||:|| |||:||   :| |  : :  :|

Db     137 GDAGCRGYYF---LRDACTYATALNVASLSVERYLAICHPFKAKTLMSRSRTKKFISAIW 193

Qy     169 GFSVLFSLP--------NTSIHGIKFHYFPNGSLV--PGSATCTVIKPMWIYNFIIQVTS 218
           | | ::|         | |  |        | ||  |   | ||         :||| :
Db     194 LASALLTVPMLFTMGEQNRSADGQH-----AGGLVCTPTIHTATV-------KVVIQVNT 241

Qy     219 FLFYLLPMTVISVLYYLMALRVSI 242
           |: :: || |||||  ::| ::::
Db     242 FMSFIFPMVVISVLNTIIANKLTV 265


RESULT 5
NTR2_RAT
ID   NTR2_RAT       STANDARD;      PRT;   416 AA.
AC   Q63384;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Neurotensin receptor type 2 (NT-R-2) (High-affinity levocabastine-
DE   sensitive neurotensin receptor).
GN   NTSR2 OR NTR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RX   MEDLINE=96228041; PubMed=8647296;
RA   Chalon P., Vita N., Kaghad M., Guillemont M., Bonin J.,
RA   Delpech B., le Fur G., Ferrara P., Caput D.;
RT   "Molecular cloning of a levocabastine-sensitive neurotensin binding
RT   site.";
RL   FEBS Lett. 386:91-94(1996).
CC   -!- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC       ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC       CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
CC       LEVELS SEEN IN THE HEART AND INTESTINE.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
CC       EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
CC       BRAIN).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       HIGHEST TO TACHYKININS RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97121; CAA65787.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003986; NT2_rec.
DR   InterPro; IPR003984; NT_rec.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01479; NEUROTENSINR.
DR   PRINTS; PR01481; NEUROTENSN2R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
FT   DOMAIN       1     32       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     33    55       1 (POTENTIAL).
FT   DOMAIN       56    64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65    87       2 (POTENTIAL).
FT   DOMAIN       88    109      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     110   131      3 (POTENTIAL).
FT   DOMAIN       132   154      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     155   176      4 (POTENTIAL).
FT   DOMAIN       177   216      EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM     217    237       5 (POTENTIAL).
FT   DOMAIN       238    297       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     298    318       6 (POTENTIAL).
FT   DOMAIN       319    337       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     338    358       7 (POTENTIAL).
FT   DOMAIN       359    416       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     108    194       BY SIMILARITY.
FT   LIPID        377    377       PALMITATE (POTENTIAL).
SQ   SEQUENCE   416 AA;  46265 MW;  127FC5F5CB6FE208 CRC64;

  Query Match             22.2%;  Score 281;  DB 1;  Length 416;
  Best Local Similarity   35.3%;  Pred. No. 1.5e-13;
  Matches   72;  Conservative   40;  Mismatches   78;  Indels   14;  Gaps    6;

      46 SVVYVPIFVVGVIGNVLVCLVILQHQAMKT-PTNYYLFSLAVSDLLVLLLGMPLEVYE-M 103
         : :|  ||  |  || |   |:|: :| :     |:: |||:| ||:||: ||:|:|  :
Db    36 TALYSLIFAFGTAGNALSVHVVLKARAGRPGRLRYHVLSLALSALLLLLVSMPMELYNFV 95

Qy   104 WRNYPFLFGPVGC---YFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA 160
         | :||::|| :||   ||   : |   :|::||: ::| || :|:  | ||:   | ||
Db    96 WSHYPWVFGDLGCRGYYF---VRELCAYATVLSVASLSAERCLAVCQPLRARRLLTPRRT 152

Qy   161 LRILGIVWGFSVLFSLPNTSIHGIKFHY-FPNGSLVPGSATCTVIKPMWIYNFIIQVTSF 219
           |:| :||  |:  :||   | | |      :|   | |  |||:         |||
Db   153 RRLLSLVWVASLGLALPMAVIMGQKHEVESADGEPEPASRVCTVLVSRATLQVFIQVNVL 212

Qy   220 LFYLLPMTVISVL-----YYLMAL 238
         : : ||: : : |      :||||
Db   213 VSFALPLALTAFLNGITVNHLMAL 236


RESULT 6
5HT1_DROME
ID   5HT1_DROME     STANDARD;      PRT;   564 AA.
AC   P20905; Q9VA21;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   5-hydroxytryptamine receptor 1 (5-HT receptor) (Serotonin receptor).
GN   5-HT7 OR 5HT-R1 OR CG12073.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oregon-R; TISSUE=Head;
RX   MEDLINE=91062395; PubMed=2174167;
RA   Witz P., Amlaiky N., Plassat J.-L., Maroteaux L., Borrelli E., Hen R.;
RT   "Cloning and characterization of a Drosophila serotonin receptor that
RT   activates adenylate cyclase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:8940-8944(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
CC       5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC       AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC       THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
CC       CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: HEAD.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       BUT WITH ONE EXTRA POTENTIAL TRANSMEMBRANE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55533; AAA28305.1; -.
DR   EMBL; AE003776; AAF57104.1; -.
DR   PIR; A38271; A38271.
DR   FlyBase; FBgn0004573; 5-HT7.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat.
FT   TRANSMEM      29     51       0 (POTENTIAL).
FT   TRANSMEM     165    188       1 (POTENTIAL).
FT   DOMAIN       189    198       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     199    222       2 (POTENTIAL).
FT   DOMAIN       223    236       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     237    258       3 (POTENTIAL).
FT   DOMAIN       259    278       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     279    302       4 (POTENTIAL).
FT   DOMAIN       303    330       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     331    353       5 (POTENTIAL).
FT   DOMAIN       354    454       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     455    476       6 (POTENTIAL).
FT   DOMAIN       477    487       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     488    510       7 (POTENTIAL).
FT   DOMAIN       511    564       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        89    106       9 X 2 AA TANDEM REPEATS OF G-S.
FT   DISULFID     235    314       BY SIMILARITY.
SQ   SEQUENCE   564 AA;  60861 MW;  0C8B9F8DA63D8095 CRC64;

  Query Match             22.2%;  Score 281;  DB 1;  Length 564;
  Best Local Similarity   36.6%;  Pred. No. 1.9e-13;
  Matches   79;  Conservative   36;  Mismatches   73;  Indels   28;  Gaps    9;
```

```
Qy      27 EEYLA--FLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSL 84
           ||: |  |:  |  | |   ||:| : : :  |:||||||: :   : :: | || | ||
Db     148 EEFAAGEFVLPPLTSIF---VSIVLLIVILGTVVGNVLVCIAVCMVRKLRRPCNYLLVSL 204

Qy      85 AVSDLLVLLLGMPL----EVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVE 140
           |:||| | || ||:    || | |      |||: |    :     | ||||:: :||:
Db     205 ALSDLCVALLVMPMALLYEVLEKWN-----FGPLLCDIWVSFDVLCCTASILNLCAISVD 259

Qy     141 RYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSAT 200
           ||:||  |    :: | || :  :|||| :   |||  | | : |   |  :
Db     260 RYLAITKPLEYGVKRTPRRMMLCVGIVWLAAACISLPPLLILGNE-HEDEEGQPI----- 313

Qy     201 CTVIKPMWIYNFIIQVTSFL--FYLLPMTVISVLYY 234
           ||| :     ||  |: : |  || :|::|:  :||
Db     314 CTVCQ-----NFAYQIYATLGSFY-IPLSVMLFVYY 343


RESULT 7
GHSR_HUMAN
ID   GHSR_HUMAN     STANDARD;      PRT;   366 AA.
AC   Q92847; Q92848; Q96RJ7;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE   peptide receptor) (GHRP) (Ghrelin receptor).
GN   GHSR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC   TISSUE=Pituitary;
RX   MEDLINE=96337998; PubMed=8688086;
RA   Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
RA   Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
RA   Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K.,
RA   McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevicz M.,
RA   Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA   Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA   Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RT   "A receptor in pituitary and hypothalamus that functions in growth
RT   hormone release.";
RL   Science 273:974-977(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RX   MEDLINE=21255649; PubMed=11356716;
RA   Petersenn S., Rasch A.C., Penshorn M., Beil F.U., Schulte H.M.;
RT   "Genomic structure and transcriptional regulation of the human growth
RT   hormone secretagogue receptor.";
RL   Endocrinology 142:2649-2659(2001).
RN   [3]
     FUNCTION.
     MEDLINE=21219832; PubMed=11322507;
     Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
RA   Tan C.P., Mckee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT   "Growth hormone secretagogue receptor family members and ligands.";
RL   Endocrine 14:9-14(2001).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=20067959; PubMed=10604470;
RA   Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT   "Ghrelin is a growth-hormone-releasing acylated peptide from
RT   stomach.";
RL   Nature 402:656-660(1999).
CC   -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC       Stimulates growth hormone secretion. Binds also other growth
CC       hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC       as well as non-peptide, low molecular weight secretagogues (e.g.
CC       L-692,429, MK-0677, adenosine).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
CC       SECRETAGOGUES.
CC   -!- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U60179; AAC50653.1; -.
DR   EMBL; U60181; AAC50654.1; -.
DR   EMBL; AF369786; AAK71539.1; -.
DR   EMBL; AF369786; AAK71540.1; -.
DR   Genew; HGNC:4267; GHSR.
DR   MIM; 601898; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN        1     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     41     66       1 (POTENTIAL).
FT   DOMAIN       67     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     96       2 (POTENTIAL).
FT   DOMAIN       97    117       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    118    139       3 (POTENTIAL).
FT   DOMAIN      140    162       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    163    183       4 (POTENTIAL).
FT   DOMAIN      184    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    212    235       5 (POTENTIAL).
FT   DOMAIN      236    263       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    264    285       6 (POTENTIAL).
FT   DOMAIN      286    302       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    303    326       7 (POTENTIAL).
FT   DOMAIN      327    366       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    116    198       BY SIMILARITY.
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    266    289       AVVVFAFILCWLPFHVGRYLFSKS -> GGSQRALRLSLAG
FT                                PILSLCLLPSL (IN ISOFORM 1B).
FT   VARSPLIC    290.   366       MISSING (IN ISOFORM 1B).
SQ   SEQUENCE   366 AA;  41328 MW;  D1B62710DA9DC0C6 CRC64;

  Query Match             21.9%;  Score 277;  DB 1;  Length 366;
  Best Local Similarity   31.7%;  Pred. No. 2.5e-13;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : ::| || || |:| ||||: |  |||::  :|
Db      46 VTATCVALFVVGIAGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 104

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     105 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 164

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::     ||:               |   ||:  :|:
Db     165 FVIWAVAFCSAGPIFVLVGVEHE---NGTDPWDTNECRPTEFAVRSGLLTVM--VWV--- 216

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
Db     217 -----SSIFFFLPVFCLTVLYSLIGRKL 239


RESULT 8
GHSR_RAT
ID   GHSR_RAT       STANDARD;      PRT;   364 AA.
```

```
AC   O08725;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE   peptide receptor) (GHRP) (Ghrelin receptor).
GN   GHSR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
     TISSUE=Pituitary;
     MEDLINE=97246555; PubMed=9092793;
     McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P.,
RA   Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
RT   "Molecular analysis of rat pituitary and hypothalamic growth hormone
RT   secretagogue receptors.";
RL   Mol. Endocrinol. 11:415-423(1997).
RN   [2]
RP   SEQUENCE OF 1-240 FROM N.A.
RC   STRAIN=Wistar; TISSUE=Pituitary;
RX   MEDLINE=98100386; PubMed=9437732;
RA   Yokote R., Sato M., Matsubara S., Ohye H., Niimi M., Murao K.,
RA   Takahara J.;
RT   "Molecular cloning and gene expression of growth hormone-releasing
RT   peptide receptor in rat tissues.";
RL   Peptides 19:15-20(1998).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=20067959; PubMed=10604470;
RA   Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT   "Ghrelin is a growth-hormone-releasing acylated peptide from
RT   stomach.";
RL   Nature 402:656-660(1999).
CC   -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC       Stimulates growth hormone secretion. Binds also other growth
CC       hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC       as well as non-peptide, low molecular weight secretagogues (e.g.
CC       L-692,429, MK-0677, adenosine) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U94321; AAC53156.1; -.
DR   EMBL; AB001982; BAA21777.1; ALT_INIT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     41     66       1 (POTENTIAL).
FT   DOMAIN       67     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     96       2 (POTENTIAL).
FT   DOMAIN       97    117       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    118    139       3 (POTENTIAL).
FT   DOMAIN      140    162       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    163    183       4 (POTENTIAL).
FT   DOMAIN      184    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    212    235       5 (POTENTIAL).
FT   DOMAIN      236    263       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    264    285       6 (POTENTIAL).
FT   DOMAIN      286    302       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    303    326       7 (POTENTIAL).
FT   DOMAIN      327    364       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    115    197       BY SIMILARITY.
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   364 AA;  40963 MW;  DCBF559BE061EEE9 CRC64;

  Query Match             21.7%;  Score 274;  DB 1;  Length 364;
  Best Local Similarity   31.7%;  Pred. No. 4.2e-13;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : ::| || || |:| ||||: |  ||||::  :|
Db      45 VTATCVALFVVGISGNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 103

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     104 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 163

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
             ::|  :   : |  : |::     ||:             |   ||:  :|:
Db     164 LVIWAVAFCSAGPIFVLVGVEHE---NGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 215

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
                 | :|: ||:  ::||| |:  ::
Db     216 -----SSVFFFLPVFCLTVLYSLIGRKL 238


RESULT 9
GHSR_PIG
ID   GHSR_PIG       STANDARD;      PRT;   366 AA.
AC   Q95254; Q95255;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE   peptide receptor) (GHRP) (Ghrelin receptor).
GN   GHSR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Yorkshire; TISSUE=Pituitary;
RX   MEDLINE=96337998; PubMed=8688086;
RA   Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
RA   Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
RA   Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K.,
RA   McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevicz M.,
RA   Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,
RA   Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA   Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RT   "A receptor in pituitary and hypothalamus that functions in growth
RT   hormone release.";
RL   Science 273:974-977(1996).
CC   -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC       Stimulates growth hormone secretion. Binds also other growth
CC       hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC       as well as non-peptide, low molecular weight secretagogues (e.g.
CC       L-692,429, MK-0677, adenosine).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
CC       SECRETAGOGUES.
CC   -!- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U60178; AAC48630.1; -.
DR   EMBL; U60180; AAC48631.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Alternative splicing.
FT   DOMAIN       1     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    41     66       1 (POTENTIAL).
FT   DOMAIN      67     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    73     96       2 (POTENTIAL).
FT   DOMAIN      97    117       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   118    139       3 (POTENTIAL).
FT   DOMAIN     140    162       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   163    183       4 (POTENTIAL).
FT   DOMAIN     184    211       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   212    235       5 (POTENTIAL).
FT   DOMAIN     236    263       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   264    285       6 (POTENTIAL).
FT   DOMAIN     286    302       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   303    326       7 (POTENTIAL).
FT   DOMAIN     327    366       CYTOPLASMIC (POTENTIAL).
FT   DISULFID   116    198       BY SIMILARITY.
FT   CARBOHYD    13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   266    289       AVVVFAFILCWLPFHVGRYLFSKS -> GGSQCALELSLPG
FT                               PLHSSCLFSSP (IN ISOFORM 1B).
FT   VARSPLIC   290    366       MISSING (IN ISOFORM 1B).
SQ   SEQUENCE   366 AA;  41194 MW;  2C850B3EF61B7C1C CRC64;

  Query Match           21.7%;  Score 274;  DB 1;  Length 366;
  Best Local Similarity 31.7%;  Pred. No. 4.2e-13;
  Matches   66;  Conservative   47;  Mismatches   69;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||: ||:|  ||: : : |:| || || |:| |||| |  |||::: :|
Db      46 VTATCVALFVVGIAGNLLTMLVVSRFREMRTTTNLYLSSMAFSDLLIFLC-MPLDLFRLW 104

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |:  | : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     105 QYRPWNLGNLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 164

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS------------LVPGSATCTVIKPMWIYNF 212
             ::|  :   : |   : |::     ||:              |   ||:  :|:
Db     165 LVIWAVAFCSAGPIFVLVGVE---HDNGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 216

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
               | :|: ||:  ::||| |:  ::
       217 -----SSVFFFLPVFCLTVLYSLIGRKL 239


RESULT 10
TRFR_CHICK
ID   TRFR_CHICK     STANDARD;      PRT;   395 AA.
AC   O93603;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
DE   receptor).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=98344812; PubMed=9681487;
RA   Sun Y.M., Millar R.P., Ho H., Gershengorn M.C., Illing N.;
RT   "Cloning and characterization of the chicken thyrotropin-releasing
RT   hormone receptor.";
RL   Endocrinology 139:3390-3398(1998).
CC   -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Y18244; CAA77091.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     30       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    31     53       1 (POTENTIAL).
FT   DOMAIN      54     63       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    64     85       2 (POTENTIAL).
FT   DOMAIN      86    101       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   102    123       3 (POTENTIAL).
FT   DOMAIN     124    146       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   147    170       4 (POTENTIAL).
FT   DOMAIN     171    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   196    217       5 (POTENTIAL).
FT   DOMAIN     218    268       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   269    290       6 (POTENTIAL).
FT   DOMAIN     291    298       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   299    321       7 (POTENTIAL).
FT   DOMAIN     322    395       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   395 AA;  44697 MW;  291D9BD2718723CB CRC64;

  Query Match           21.4%;  Score 270.5;  DB 1;  Length 395;
  Best Local Similarity 32.0%;  Pred. No. 8e-13;
  Matches   72;  Conservative   52;  Mismatches   78;  Indels   23;  Gaps    6;

Qy      20 QKH---LNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           | |   | |::|::        : : :   :: : |  :|::||::| ||:|: : |:||
Db       9 QNHTGLLLSSQEFVT-------AEYQVVTILLVLLICGLGIVGNIMVVLVVLRTKHMRTP 61

Qy      77 TNYYLFSLAVSDLLVLL-LGMP---LEVYEMWRNYPFLFGPVGCYFKTALFETVCFASIL 132
           || || ||||:||:||:  |:|     :|: |     ::| |||   | |      ||
Db      62 TNCYLVSLAVADLMVLVAAGLPNITESLYKSW-----VYGYVGCLCITYLQYLGINASSF 116

Qy     133 SITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNG 192
           |||  ::|||:|| || :|:   |  || :|:  || |: :: :    :  :    : :
Db     117 SITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWSFASVYCMLWFFLLDLNIAVYKDT 176

Qy     193 SLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA 237
           ::|    :|       |: |  :   :||:||| : :||| |:|
Db     177 TVV----SCGYKVSRSYYSPIYMMDFGIFYVLPMVLATVLYGLIA 217


RESULT 11
GHSR_MOUSE
ID   GHSR_MOUSE     STANDARD;      PRT;   257 AA.
AC   Q99P50; Q91Z82;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
```

```
DE   Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
DE   peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).
GN   GHSR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE OF 1-183 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RA   Kacsoh B.;
RT   "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by
RT   rapid amplification of cDNA ends (RACE).";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 73-257 FROM N.A.
RC   STRAIN=129S3/SvImJ;
RA   Peng X., Frohman L.A., Kineman R.D.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
CC       Stimulates growth hormone secretion. Binds also other growth
CC       hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
CC       as well as non-peptide, low molecular weight secretagogues (e.g.
CC       L-692,429, MK-0677, adenosine) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY056474; AAL13336.1; -.
DR   EMBL; AF332997; AAG61141.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     35       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     36    58       1 (POTENTIAL).
FT   DOMAIN       59    77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78    100      2 (POTENTIAL).
FT   DOMAIN       101   125      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     126   148      3 (POTENTIAL).
FT   DOMAIN       149   160      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     161   183      4 (POTENTIAL).
FT   DOMAIN       184   211      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     212   234      5 (POTENTIAL).
FT   DOMAIN       235   257      CYTOPLASMIC (POTENTIAL).
FT   DISULFID     115   197      BY SIMILARITY.
FT   CARBOHYD     13    13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26    26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     187   187      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER      257   257
SQ   SEQUENCE   257 AA;  28743 MW;  D6FA21EAB0E30791 CRC64;

  Query Match             21.2%;  Score 268;  DB 1;  Length 257;
  Best Local Similarity   31.2%;  Pred. No. 8.2e-13;
  Matches   65;  Conservative   47;  Mismatches   70;  Indels   26;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           |:   | :||||:  |:|  ||: : : ::| || || |:| ||||: |  |||::  :|
Db      45 VTATCVALFVVGISSNLLTMLVVSRFRELRTTTNLYLSSMAFSDLLIFLC-MPLDLVRLW 103

Qy     105 RNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRIL 164
           :  |: || : |     : |:  :|::|:|| :||||| ||  | |||:  |: |   ::
Db     104 QYRPWNFGDLLCKLFQFVSESCTYATVLTITALSVERYFAICFPLRAKVVVTKGRVKLVI 163

Qy     165 GIVWGFSVLFSLPNTSIHGIKFHYFPNGS-----------LVPGSATCTVIKPMWIYNF 212
           ::|  :   : |   : |::     ||:                |   ||:  :|:
Db     164 LVIWAVAFCSAGPIFVLVGVEHE---NGTDPRDTNECRATEFAVRSGLLTVM--VWV--- 215

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
                 | :|: ||:  ::||| |:  ::
Db     216 -----SSVFFFLPVFCLTVLYSLIGRKL 238


RESULT 12
TRFR_MOUSE
ID   TRFR_MOUSE     STANDARD;      PRT;   393 AA.
AC   P21761;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
DE   receptor).
GN   TRHR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=91088548; PubMed=2175902;
RA   Straub R.E., Frech G.C., Joho R.H., Gershengorn M.C.;
RT   "Expression cloning of a cDNA encoding the mouse pituitary
RT   thyrotropin-releasing hormone receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:9514-9518(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92381047; PubMed=1324930;
RA   Narayanan C.S., Fujimoto J., Geras-Raaka E., Gershengorn M.C.;
RT   "Regulation by thyrotropin-releasing hormone (TRH) of TRH receptor
RT   mRNA degradation in rat pituitary GH3 cells.";
RL   J. Biol. Chem. 267:17296-17303(1992).
RN   [3]
RP   SEQUENCE OF 332-393 FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=97013702; PubMed=9156522;
RA   Jones K.E., Brubaker J.H., Chin W.W.;
RT   "An alternative splice variant of the mouse TRH receptor mRNA is the
RT   major form expressed in the mouse pituitary gland.";
RL   J. Mol. Endocrinol. 16:197-204(1996).
CC   -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M59811; AAA40480.1; -.
DR   EMBL; M94384; AAA40437.1; -.
DR   EMBL; L48936; AAA81559.1; -.
DR   PIR; A39251; A39251.
DR   MGD; MGI:98824; Trhr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29    51       1 (POTENTIAL).
FT   DOMAIN       52    61       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     62     83       2 (POTENTIAL).
FT   DOMAIN       84     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    100    121       3 (POTENTIAL).
FT   DOMAIN      122    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    168       4 (POTENTIAL).
FT   DOMAIN      169    193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    194    215       5 (POTENTIAL).
FT   DOMAIN      216    266       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    267    288       6 (POTENTIAL).
FT   DOMAIN      289    296       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    297    319       7 (POTENTIAL).
FT   DOMAIN      320    393       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   393 AA;  44559 MW;  8739B75D1A0FCCFC CRC64;

  Query Match             21.0%;  Score 265;  DB 1;  Length 393;
  Best Local Similarity   28.9%;  Pred. No. 2e-12;
  Matches   72;  Conservative   48;  Mismatches   69;  Indels   60;  Gaps    9;

Qy      15 LEDPFQKHLNSTE---------EY-------LAFLCGPRRSHFFLPVSVVYVPIFVVGVI 58
           :|:     :| ||         ||       :  :||                   :|::
Db       1 MENDTVSEMNQTELQPQAAVALEYQVVTILLVVIICG------------------LGIV 41

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---EVYEMWRNYPFLFGPV 114
           ||::| ||::: : |:|||| || ||||:||:||:  |:|     :|  |     ::| |
Db      42 GNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDSIYGSW-----VYGYV 96

Qy     115 GCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLF 174
           ||   | |      ||  |||  ::|||:|| || :|:   |  || :|:  || |: ::
Db      97 GCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSRAKKIIIFVWAFTSIY 156

Qy     175 SLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFIIQVTSFLFYLLPMTV 228
           :   :  :   : |  :|    :|         |::: :| :      ||::|| :
Db     157 CMLWFFLLDLNISTYKNAVVV----SCGYKISRNYYSPIYLMDFGV------FYVVPMIL 206

Qy     229 ISVLYYLMA 237
           :|||  :|
Db     207 ATVLYGFIA 215


RESULT 13
TRFR_HUMAN
ID   TRFR_HUMAN     STANDARD;      PRT;   398 AA.
AC   P34981;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
DE   receptor).
GN   TRHR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
     NCBI_TaxID=9606;
     [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93371401; PubMed=8395824;
RA   Matre V., Karlsen H.E., Wright M.S., Lundell I., Fjeldheim K.,
RA   Gabrielsen O.S., Larhammar D., Gautvik K.M.;
RT   "Molecular cloning of a functional human thyrotropin-releasing
RT   hormone receptor.";
RL   Biochem. Biophys. Res. Commun. 195:179-185(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=93384596; PubMed=8396925;
RA   Yamada M., Monden T., Satoh T., Satoh N., Murakami M., Iriuchijima T.,
RA   Kakegawa T., Mori M.;
RT   "Pituitary adenomas of patients with acromegaly express thyrotropin-
RT   releasing hormone receptor messenger RNA: cloning and functional
RT   expression of the human thyrotropin-releasing hormone receptor
RT   gene.";
RL   Biochem. Biophys. Res. Commun. 195:737-745(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94063224; PubMed=8243797;
RA   Duthie S.M., Taylor P.L., Anderson L., Cook J., Eidne K.A.;
RT   "Cloning and functional characterisation of the human TRH receptor.";
RL   Mol. Cell. Endocrinol. 95:R11-R15(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95002135; PubMed=7918619;
RA   Hinuma S., Hosoya M., Ogi K., Tanaka H., Nagai Y., Onda H.;
RT   "Molecular cloning and functional expression of a human thyrotropin-
RT   releasing hormone (TRH) receptor gene.";
RL   Biochim. Biophys. Acta 1219:251-259(1994).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96355621; PubMed=8703031;
RA   Iwasaki T., Yamada M., Satoh T., Konaka S., Ren Y., Hashimoto K.,
RA   Kohga H., Kato Y., Mori M.;
RT   "Genomic organization and promoter function of the human thyrotropin-
RT   releasing hormone receptor gene.";
RL   J. Biol. Chem. 271:22183-22188(1996).
RN   [6]
RP   SEQUENCE OF 1-263 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=99101143; PubMed=9886052;
RA   Matre V., Hovring P.I., Orstavik S., Frengen E., Rian E.,
RA   Velickovic Z., Murray-Mcintosh R.P., Gautvik K.M.;
RT   "Structural and functional organization of the gene encoding the
RT   human thyrotropin-releasing hormone receptor.";
RL   J. Neurochem. 72:40-50(1999).
CC   -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16845; BAA04120.1; -.
DR   EMBL; X75071; CAA52965.1; -.
DR   EMBL; X72089; CAA50979.1; -.
DR   EMBL; S75283; AAB32222.1; -.
DR   EMBL; S75281; AAB32222.1; JOINED.
DR   EMBL; D85376; BAA12796.1; -.
DR   EMBL; D85375; BAA12796.1; JOINED.
DR   EMBL; AJ011701; CAA09746.1; -.
DR   PIR; JN0759; JN0759.
DR   PIR; S40682; S40682.
DR   PIR; JN0708; JN0708.
DR   Genew; HGNC:12299; TRHR.
DR   MIM; 188545; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT   DOMAIN       1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    29     51       1 (POTENTIAL).
FT   DOMAIN      52     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    62     83       2 (POTENTIAL).
FT   DOMAIN      84     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   100    121       3 (POTENTIAL).
FT   DOMAIN     122    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   145    168       4 (POTENTIAL).
FT   DOMAIN     169    193       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM     194    215       5 (POTENTIAL).
FT   DOMAIN       216    266       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     267    288       6 (POTENTIAL).
FT   DOMAIN       289    296       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     297    319       7 (POTENTIAL).
FT   DOMAIN       320    398       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT       10     10       N -> K (IN DBSNP:5774).
FT                                 /FTId=VAR_011857.
SQ   SEQUENCE   398 AA;  45084 MW;  FE920B5F2E933D3E CRC64;

  Query Match             20.5%;  Score 259;  DB 1;  Length 398;
  Best Local Similarity   32.4%;  Pred. No. 5.4e-12;
  Matches   66;  Conservative   47;  Mismatches   65;  Indels   26;  Gaps    7;

Qy      45 VSVVYVPIFV-VGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---E 99
           |::: | |   :|::||::| ||::: : |:|||| || ||||:||:||:  |:|
Db      27 VTILLVLIICGLGIVGNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDS 86

Qy     100 VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRR 159
           :|  |     ::| |||   | |      || |||  ::|||:|| || :|:   |  |
Db      87 IYGSW-----VYGYVGCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSR 141

Qy     160 ALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFI 213
           | :|:  || |: |: :    :  :    : :  ::    :|          |::: :|
Db     142 AKKIIIFVWAFTSLYCMLWFFLLDLNISTYKDAIVI----SCGYKISRNYYSPIYLMDFG 197

Qy     214 IQVTSFLFYLLPMTVISVLYYLMA 237
           :      ||::|| : :|||  :|
Db     198 V------FYVVPMILATVLYGFIA 215


RESULT 14
TRFR_BOVIN
ID   TRFR_BOVIN     STANDARD;      PRT;   398 AA.
AC   O46639;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
DE   receptor).
GN   TRHR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RX   MEDLINE=98151337; PubMed=9492373;
RA   Takata M., Shimada Y., Ikeda A., Sekikawa K.;
RT   "Molecular cloning of bovine thyrotropin-releasing hormone receptor
RT   gene.";
RL   J. Vet. Med. Sci. 60:123-127(1998).
CC   -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D83964; BAA24069.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN         1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      29     51       1 (POTENTIAL).
FT   DOMAIN        52     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      62     83       2 (POTENTIAL).
FT   DOMAIN        84     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     100    121       3 (POTENTIAL).
FT   DOMAIN       122    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     145    168       4 (POTENTIAL).
FT   DOMAIN       169    193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     194    215       5 (POTENTIAL).
FT   DOMAIN       216    266       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     267    288       6 (POTENTIAL).
FT   DOMAIN       289    296       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     297    319       7 (POTENTIAL).
FT   DOMAIN       320    398       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   398 AA;  45165 MW;  D9AF4B211A5701B8 CRC64;

  Query Match             20.4%;  Score 258;  DB 1;  Length 398;
  Best Local Similarity   32.4%;  Pred. No. 6.3e-12;
  Matches   66;  Conservative   47;  Mismatches   65;  Indels   26;  Gaps    7;

Qy      45 VSVVYVPIFV-VGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---E 99
           |::: | |   :|::||::| ||::: : |:|||| || ||||:||:||:  |:|
Db      27 VTILLVLIICGLGIVGNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDS 86

Qy     100 VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRR 159
           :|  |     ::| |||   | |      || |||  ::|||:|| || :|:   |  |
Db      87 IYGSW-----VYGYVGCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSR 141

Qy     160 ALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFI 213
           | :|:  || |: :: :    :  :    : :  :|    :|          |::: :|
Db     142 AKKIIIFVWAFTSIYCMLWFFLLDLNISTYKDAIVV----SCGYKISRNYYSPIYLMDFG 197

Qy     214 IQVTSFLFYLLPMTVISVLYYLMA 237
           :      ||::|| : :|||  :|
Db     198 V------FYVVPMILATVLYGFIA 215


RESULT 15
TRFR_SHEEP
ID   TRFR_SHEEP     STANDARD;      PRT;   398 AA.
AC   Q28596;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
DE   receptor).
GN   TRHR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=97200775; PubMed=9048604;
RA   Bockmann J., Boeckers T.M., Winter C., Wittkowski W., Winterhoff H.,
RA   Deufel T., Kreutz M.R.;
RT   "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
RT   is 3,5,3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1
RT   independent.";
RL   Endocrinology 138:1019-1028(1997).
CC   -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC       PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X95285; CAA64606.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     28       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     29     51       1 (POTENTIAL).
FT   DOMAIN       52     61       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     62     83       2 (POTENTIAL).
FT   DOMAIN       84     99       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    100    121       3 (POTENTIAL).
FT   DOMAIN      122    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    168       4 (POTENTIAL).
FT   DOMAIN      169    193       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    194    215       5 (POTENTIAL).
FT   DOMAIN      216    266       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    267    288       6 (POTENTIAL).
FT   DOMAIN      289    296       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    297    319       7 (POTENTIAL).
FT   DOMAIN      320    398       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     10     10       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   398 AA;  45088 MW;  375A311D3DD2A61A CRC64;

  Query Match             20.4%;  Score 258;  DB 1;  Length 398;
  Best Local Similarity   32.4%;  Pred. No. 6.3e-12;
  Matches   66;  Conservative   47;  Mismatches   65;  Indels   26;  Gaps    7;

Qy      45 VSVVYVPIFV-VGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLL-LGMPL---E 99
           |::: | |   :|::||::| ||::: : |:|||| || ||||:||:||:  |:|
Db      27 VTILLVLIICGLGIVGNIMVVLVVMRTKHMRTPTNCYLVSLAVADLMVLVAAGLPNITDS 86

Qy     100 VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRR 159
           :|  |     ::| |||   | |      ||  |||  ::|||:|| || :|:   |  |
Db      87 IYGSW-----VYGYVGCLCITYLQYLGINASSCSITAFTIERYIAICHPIKAQFLCTFSR 141

Qy     160 ALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATC------TVIKPMWIYNFI 213
           | :|:  || |: :: :    :  :    : :  :|    :|          |::: :|
Db     142 AKKIIIFVWAFTSIYCMLWFFLLDLNISTYKDAIVV----SCGYKISRNYYSPIYLMDFG 197

Qy     214 IQVTSFLFYLLPMTVISVLYYLMA 237
           :      ||::|| : :|||  :|
Db     198 V------FYVVPMILATVLYGFIA 215


Search completed: January 17, 2003, 05:04:48
Job time : 35 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:04:56 ; Search time 64 Seconds
(without alignments)
75.166 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	418	33.1	419	10	US-09-804-551B-26	Sequence 26, Appl
2	262.5	20.8	222	10	US-09-911-005-4	Sequence 4, Appli
3	262.5	20.8	259	9	US-10-080-960-25	Sequence 25, Appl
4	262.5	20.8	259	9	US-10-012-140-27	Sequence 27, Appl
5	262.5	20.8	259	9	US-10-012-140-29	Sequence 29, Appl
6	262.5	20.8	259	9	US-10-077-698-3	Sequence 3, Appli
7	262.5	20.8	259	10	US-09-796-338A-23	Sequence 23, Appl
8	256.5	20.3	415	10	US-09-823-114-20	Sequence 20, Appl
9	255	20.2	390	10	US-09-761-962-25	Sequence 25, Appl
10	255	20.2	391	10	US-09-761-962-26	Sequence 26, Appl
11	255	20.2	392	10	US-09-761-962-19	Sequence 19, Appl
12	255	20.2	398	10	US-09-214-904-2	Sequence 2, Appli
13	255	20.2	398	10	US-09-761-962-29	Sequence 29, Appl
14	255	20.2	398	10	US-09-966-871-83	Sequence 83, Appl
15	255	20.2	398	12	US-10-039-645-83	Sequence 83, Appl
16	255	20.2	401	10	US-09-761-962-20	Sequence 20, Appl
17	255	20.2	409	10	US-09-761-962-27	Sequence 27, Appl
18	255	20.2	438	10	US-09-761-962-17	Sequence 17, Appl
19	255	20.2	444	10	US-09-761-962-28	Sequence 28, Appl
20	254.5	20.2	405	10	US-09-966-871-84	Sequence 84, Appl
21	254.5	20.2	405	12	US-10-039-645-84	Sequence 84, Appl
22	253.5	20.1	400	10	US-09-966-871-85	Sequence 85, Appl
23	253.5	20.1	400	12	US-10-039-645-85	Sequence 85, Appl
24	253	20.0	383	10	US-09-966-871-87	Sequence 87, Appl
25	253	20.0	383	12	US-10-039-645-87	Sequence 87, Appl
26	251.5	19.9	400	10	US-09-966-871-86	Sequence 86, Appl
27	251.5	19.9	400	12	US-10-039-645-86	Sequence 86, Appl
28	247.5	19.6	369	10	US-09-823-114-9	Sequence 9, Appli
29	246.5	19.5	400	10	US-09-966-871-78	Sequence 78, Appl
30	246.5	19.5	400	12	US-10-039-645-78	Sequence 78, Appl
31	244.5	19.4	506	10	US-09-804-551B-34	Sequence 34, Appl
32	242.5	19.2	367	10	US-09-823-114-23	Sequence 23, Appl
33	242.5	19.2	382	10	US-09-993-844-4	Sequence 4, Appli
34	242.5	19.2	398	10	US-09-823-114-16	Sequence 16, Appl
35	242.5	19.2	398	10	US-09-966-871-1	Sequence 1, Appli
36	242.5	19.2	398	12	US-10-039-645-1	Sequence 1, Appli
37	242	19.2	359	10	US-09-966-871-81	Sequence 81, Appl
38	242	19.2	359	12	US-10-039-645-81	Sequence 81, Appl
39	240.5	19.0	398	10	US-09-966-871-79	Sequence 79, Appl
40	240.5	19.0	398	12	US-10-039-645-79	Sequence 79, Appl
41	239	18.9	359	9	US-09-867-915-3	Sequence 3, Appli
42	239	18.9	359	10	US-09-867-569-9	Sequence 9, Appli
43	238.5	18.9	370	10	US-09-823-114-19	Sequence 19, Appl
44	237	18.8	460	9	US-10-166-101-4	Sequence 4, Appli
45	234.5	18.6	453	9	US-09-970-966-209	Sequence 209, App

ALIGNMENTS

RESULT 1
US-09-804-551B-26
; Sequence 26, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804,551B
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-26

Query Match 33.1%; Score 418; DB 10; Length 419;
Best Local Similarity 40.9%; Pred. No. 1e-32;
Matches 85; Conservative 49; Mismatches 62; Indels 12; Gaps 4;

```
Qy      35 GPRRS--HFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVL 92
           || |      :||:|||  ||: ||:||:  |:|| ::::| | |||||||||:|| |:|
Db      11 GPPRDPLAIVIPVTVVYSLIFITGVVGNISTCIVIKKNRSMHTATNYYLFSLAISDFLLL 70

Qy      93 LLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAK 152
           | |:| ||  :|  ||::||   |  :  | ||   |::|:||  :||||:|| |||  :
Db      71 LSGVPQEVSYIWSKYPYVFGEYICIGRGLLAETSANATVLTITAFTVERYIAICHPFLGQ 130

Qy     153 LQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNF 212
             |   ||:||: :||  ::: ::|  :  ||: ||        |   | ::: :  ::|
Db     131 AMSKLSRAIRIIVLVWIMAIVTAIPQAAQFGIE-HY-------SGVEQCGIVRVIVKHSF 182

Qy     213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
           |:::|:|:| ||::| ||| |: : :
Db     183 --QLSTFIFFLAPMSIILVLYLLIGVHL 208

RESULT 2
```

```
US-09-911-005-4
; Sequence 4, Application US/09911005
; Patent No. US20020119493A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  TITLE OF INVENTION: 65494, A NOVEL HUMAN G-PROTEIN-COUPLED
;  TITLE OF INVENTION:  RECEPTOR FAMILY MEMBER AND USES THEREOF
;  FILE REFERENCE: 10448-076001
;  CURRENT APPLICATION NUMBER: US/09/911,005
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: 60/220,042
;  PRIOR FILING DATE: 2000-07-21
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 222
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence
US-09-911-005-4

  Query Match             20.8%;  Score 262.5;  DB 10;  Length 222;
  Best Local Similarity   35.3%;  Pred. No. 4e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||| : ::|||| :: :|||:||| ||   |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
           |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||      :|     ||:|      | :  |         |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 3
US-10-080-960-25
; Sequence 25, Application US/10080960
; Publication No. US20020197695A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Glucksmann, Maria
;  APPLICANT:  Meyers, Rachel
;  TITLE OF INVENTION: 80090, 52874,52880,63497, AND 33425
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF
;  FILE REFERENCE: 38155-20044.00
;  CURRENT APPLICATION NUMBER: US/10/080,960
;  CURRENT FILING DATE:  2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/242,040
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/242,038
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/241,992
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: US 60/242,637
;  PRIOR FILING DATE: 2000-10-23
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Consensus amino acid sequence
US-10-080-960-25

  Query Match             20.8%;  Score 262.5;  DB 9;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 4.7e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||| : ::|||| :: :|||:||| ||   |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
           |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||      :|     ||:|      | :  |         |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 4
US-10-012-140-27
; Sequence 27, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
;  APPLICANT: Leiby, Kevin R.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
;  TITLE OF INVENTION:  14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 381552004900
;  CURRENT APPLICATION NUMBER: US/10/012,140
;  CURRENT FILING DATE:  2001-11-08
;  PRIOR APPLICATION NUMBER: 60/246,768
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/246,772
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/249,185
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Consensus amino acid sequence
US-10-012-140-27

  Query Match             20.8%;  Score 262.5;  DB 9;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 4.7e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||| : ::|||| :: :|||:||| ||   |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
           |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||      :|     ||:|      | :  |         |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184
```

```
RESULT 5
US-10-012-140-29
; Sequence 29, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
;  APPLICANT: Leiby, Kevin R.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
;  TITLE OF INVENTION:  14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 381552004900
;  CURRENT APPLICATION NUMBER: US/10/012,140
;  CURRENT FILING DATE:  2001-11-08
   PRIOR APPLICATION NUMBER: 60/246,768
   PRIOR FILING DATE: 2000-11-08
,  PRIOR APPLICATION NUMBER: 60/246,772
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/249,185
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Consensus amino acid sequence
US-10-012-140-29

  Query Match             20.8%;  Score 262.5;  DB 9;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 4.7e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||||: : ::|||| :: :|||:||| ||    |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
             |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||       :|     ||:|      | :  |          |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 6
US-10-077-698-3
; Sequence 3, Application US/10077698
; Publication No. US20030008350A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: 14273 Receptor, A No. US20030008350A1el G-Protein Coupled Recepto
;  FILE REFERENCE: 5800-4B, 035800/177086
;  CURRENT APPLICATION NUMBER: US/10/077,698
;  CURRENT FILING DATE:  2002-02-13
;  PRIOR APPLICATION NUMBER: 09/261,599
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 09/107,761
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: 09/223,538
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Seven
;   OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily
US-10-077-698-3

  Query Match             20.8%;  Score 262.5;  DB 9;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 4.7e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||||: : ::|||| :: :|||:||| ||    |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
             |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||       :|     ||:|      | :  |          |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 7
US-09-796-338A-23
; Sequence 23, Application US/09796338A
; Patent No. US20020061522A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
;  TITLE OF INVENTION:  52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
;  FILE REFERENCE: 10448-020001
;  CURRENT APPLICATION NUMBER: US/09/796,338A
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,059
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence
US-09-796-338A-23

  Query Match             20.8%;  Score 262.5;  DB 10;  Length 259;
  Best Local Similarity   35.3%;  Pred. No. 4.7e-18;
  Matches   67;  Conservative   38;  Mismatches   64;  Indels   21;  Gaps    5;

Qy      59 GNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVY------EMWRNYPFLFG 112
           ||:|| |||||: : ::|||| :: :|||:||| ||    |  :|      | |      ||
Db       1 GNLLVILVILRTKKLRTPTNIFILNLAVADLLFLLTLPPWALYYLVGGSEDWP-----FG 55

Qy     113 PVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR-RRALRILGIVWGFS 171
             |   |||     :|||| :| :|::||:||:|| | : : |  |||  :: :||  :
Db      56 SALCKLVTALDVVNMYASILLLTAISIDRYLAIVHPLRYRRRRTSPRRAKVVILLVWVLA 115

Qy     172 VLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKP--------MWIYNFIIQVTSFLFYL 223
           :| |||       :|     ||:|      | :  |          |: :::: ::: : :|
Db     116 LLLSLPPLLFSWVKTVEEGNGTLNVNVTVCLIDFPEESTASVSTWLRSYVL-LSTLVGFL 174

Qy     224 LPMTVISVLY 233
           ||: || | |
Db     175 LPLLVILVCY 184


RESULT 8
US-09-823-114-20
; Sequence 20, Application US/09823114
```

```
; Patent No. US20020061554A1
;  GENERAL INFORMATION:
;       APPLICANT: EVANS, CHRISTOPHER J.
;                  KEITH, DUANE E.
;       TITLE OF INVENTION: OPIOID RECEPTOR GENES
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: MORRISON & FOERSTER
;            STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
;            CITY: WASHINGTON
;            STATE: DC
;            COUNTRY: USA
;            ZIP: 20006-1888
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/823,114
;            FILING DATE: 29-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/148,351
:            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: MURASHIGE, KATE H.
;            REGISTRATION NUMBER: 29,959
;            REFERENCE/DOCKET NUMBER: 22000-20526.22
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 887-1500
;            TELEFAX: (202) 887-0763
;            TELEX: 90-4030 MRSNFOERSWSH
;   INFORMATION FOR SEQ ID NO: 20:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 415 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  group(9, 12, 33, 40, 48)
;            OTHER INFORMATION:  /note= "extracellular Asn residues
;            that are consensus sites for N-linked glycosylation"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-823-114-20

  Query Match             20.3%;  Score 256.5;  DB 10;  Length 415;
  Best Local Similarity   28.3%;  Pred. No. 3e-17;
  Matches   72;  Conservative   49;  Mismatches   90;  Indels   43;  Gaps    8;

Qy       8 SWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF----------------FLPVSVVYV 50
           ||:           :|:  :  |:  ||| |::                : :  :|
Db      30 SWV----------NLSHLDGNLSDPCGPNRTNLGGRDSLCPPTGSPSMITAITIMALYS 78

Qy      51 PIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPF 109
             : |||: || ||  ||:::  ||| || |:|:||::| |     :| : |  :   :|
Db      79 IVCVVGLFGNFLVMYVIVRYTKMKTATNIYIFNLALADALATST-LPFQSVNYLMGTWP- 136

Qy     110 LFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWG 169
             || : |    ::      | || :: |:||:||:|: || :|    | | |  |     |
Db     137 -FGTILCKIVISIDYYNMFTSIFTLCTMSVDRYIAVCHPVKALDFRTPRNAKIINVCNWI 195

Qy     170 FSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPM 226
             |    ||   :   |:          ||  ||:    | | :  ::::  |:| :::|:
Db     196 LSSAIGLPVMFMATTKYR--------QGSIDCTLTFSHPTWYWENLVKICVFIFAFIMPV 247

Qy     227 TVISVLYYLMALRV 240
             :|:| | || ||:
Db     248 LIITVCYGLMILRL 261


RESULT 9
```

```
US-09-761-962-25
; Sequence 25, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: Identification and Characterization of Multiple Splice
;  TITLE OF INVENTION:  Variants of Mu-
;  TITLE OF INVENTION:  opioid Receptor (MOR-1) Gene
;  FILE REFERENCE: 830002-2000.1
;  CURRENT APPLICATION NUMBER: US/09/761,962
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 09/743,872
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-761-962-25

  Query Match             20.2%;  Score 255;  DB 10;  Length 390;
  Best Local Similarity   30.7%;  Pred. No. 3.9e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP-----------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           |||  |     ||   |           :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db     212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 10
US-09-761-962-26
; Sequence 26, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: Identification and Characterization of Multiple Splice
;  TITLE OF INVENTION:  Variants of Mu-
;  TITLE OF INVENTION:  opioid Receptor (MOR-1) Gene
;  FILE REFERENCE: 830002-2000.1
;  CURRENT APPLICATION NUMBER: US/09/761,962
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 09/743,872
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
;   LENGTH: 391
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-761-962-26

  Query Match             20.2%;  Score 255;  DB 10;  Length 391;
  Best Local Similarity   30.7%;  Pred. No. 3.9e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP-----------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           |||  |     ||   |           :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
```

```
          || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db    103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy    136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
          |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db    160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy    196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
            ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db    212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 11
US-09-761-962-19
; Sequence 19, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: Identification and Characterization of Multiple Splice
;  TITLE OF INVENTION:  Variants of Mu-
;  TITLE OF INVENTION:  opioid Receptor (MOR-1) Gene
;  FILE REFERENCE: 830002-2000.1
;  CURRENT APPLICATION NUMBER: US/09/761,962
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 09/743,872
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
;   LENGTH: 392
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-761-962-19

  Query Match             20.2%;  Score 255;  DB 10;  Length 392;
  Best Local Similarity   30.7%;  Pred. No. 3.9e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP------------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           ||| |     ||   |             :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db     212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 12
US-09-214-904-2
; Sequence 2, Application US/09214904
; Patent No. US20010047519A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  TRANSGENIC ANIMAL IN WHICH THE EXPRESSION
;    TITLE OF INVENTION:  OF OPIATE RECEPTORS IS MODIFIED
;    NUMBER OF SEQUENCES:  6
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/214,904
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR97/01282
;      FILING DATE:
;      APPLICATION NUMBER:  FR 96.08810
;      FILING DATE:  15-JUL-1996
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  398 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-214-904-2

  Query Match             20.2%;  Score 255;  DB 10;  Length 398;
  Best Local Similarity   30.7%;  Pred. No. 4e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP------------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           ||| |     ||   |             :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db     212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 13
US-09-761-962-29
; Sequence 29, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: Identification and Characterization of Multiple Splice
;  TITLE OF INVENTION:  Variants of Mu-
;  TITLE OF INVENTION:  opioid Receptor (MOR-1) Gene
;  FILE REFERENCE: 830002-2000.1
;  CURRENT APPLICATION NUMBER: US/09/761,962
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 09/743,872
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-761-962-29

  Query Match             20.2%;  Score 255;  DB 10;  Length 398;
  Best Local Similarity   30.7%;  Pred. No. 4e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP------------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           ||| |     ||   |             :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
```

```
Db    212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 14
US-09-966-871-83
; Sequence 83, Application US/09966871
; Patent No. US20020127539A1
; GENERAL INFORMATION:
;  APPLICANT: Kopin, Alan S.
;  TITLE OF INVENTION: Assays for Identifying Receptors Having
;  TITLE OF INVENTION:  Alterations in Signaling
;  FILE REFERENCE: 00398/512002
;  CURRENT APPLICATION NUMBER: US/09/966,871
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: US 60/236,302
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/288,644
;  PRIOR FILING DATE: 2001-05-03
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-966-871-83

  Query Match            20.2%;  Score 255;  DB 10;  Length 398;
  Best Local Similarity  30.7%;  Pred. No. 4e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP-----------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           ||| |     ||   |            :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db     212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


RESULT 15
US-10-039-645-83
; Sequence 83, Application US/10039645
; Patent No. US20020147170A1
; GENERAL INFORMATION:
;  APPLICANT: Kopin, Alan S.
   APPLICANT:  Beinborn, Martin
   TITLE OF INVENTION: Constitutively Active, Hypersensitive,
   TITLE OF INVENTION:  and No. US20020147170A1functional Receptors as No. US20020147170
;  FILE REFERENCE: 00398/510002
;  CURRENT APPLICATION NUMBER: US/10/039,645
;  CURRENT FILING DATE:  2001-10-25
;  PRIOR APPLICATION NUMBER: US 60/243,550
;  PRIOR FILING DATE: 2000-10-26
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-039-645-83

  Query Match            20.2%;  Score 255;  DB 12;  Length 398;
  Best Local Similarity  30.7%;  Pred. No. 4e-17;
  Matches   70;  Conservative   42;  Mismatches   84;  Indels   32;  Gaps    8;

Qy      34 CGPRR-----SHFFLP-----------VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTP 76
           ||| |     ||   |            :  :|  : |||: || ||  ||:::  |||
Db      43 CGPNRTGLGGSHSLCPQTGSPSMVTAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKTA 102

Qy      77 TNYYLFSLAVSDLLVLLLGMPLE-VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           || |:|:||::| |     :| : |  :   :|  || : |    ::     | || ::
Db     103 TNIYIFNLALADALATST-LPFQSVNYLMGTWP--FGNILCKIVISIDYYNMFTSIFTLC 159

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           |:||:||:|: || :|    | | |  :    |  |    ||   :   |:
Db     160 TMSVDRYIAVCHPVKALDFRTPRNAKIVNVCNWILSSAIGLPVMFMATTKYR-------- 211

Qy     196 PGSATCTVI--KPMWIYNFIIQVTSFLF-YLLPMTVISVLYYLMALRV 240
             ||  ||:    | | :  ::::  |:| :::|: :|:| | || ||:
Db     212 QGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLMILRL 259


Search completed: January 17, 2003, 05:09:27
Job time : 67 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:35:21 ; Search time 2200 Seconds
(without alignments)
5366.595 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc.....cactcagagtgagtatctag 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	528	72.4	540	14	BM737726	BM737726 K-EST0000
c	2	488.4	67.0	670	17	AG046172	AG046172 Pan trogl
	3	354.2	48.6	555	12	BE751626	BE751626 203809 MA
	4	315	43.2	929	12	BE871983	BE871983 601448032
	5	280	38.4	624	10	BB587450	BB587450 BB587450
	6	275.4	37.8	465	12	BE751310	BE751310 203132 MA
	7	230.2	31.6	993	14	BQ721315	BQ721315 AGENCOURT
c	8	211.2	29.0	941	17	CNS036WH	AL230570 Tetraodon
	9	200.6	27.5	624	9	AU244292	AU244292 AU244292
	10	166.6	22.9	445	9	AU244285	AU244285 AU244285
	11	116.4	16.0	959	17	CNS04GXP	AL290230 Tetraodon
c	12	107.8	14.8	817	17	CNS01XJ2	AL171767 Tetraodon
	13	107	14.7	500	12	BF603623	BF603623 269181 MA
	14	97.6	13.4	1027	13	BM543497	BM543497 AGENCOURT
	15	95.4	13.1	448	12	BF323227	BF323227 maa38a07.
	16	95.2	13.1	975	17	CNS039RV	AL234292 Tetraodon
c	17	95	13.0	835	17	CNS02QN1	AL209494 Tetraodon
	18	93.4	12.8	1010	14	BM925480	BM925480 AGENCOURT
	19	92.4	12.7	855	13	BI757121	BI757121 603030882
	20	91.8	12.6	1055	14	BM808829	BM808829 AGENCOURT
	21	91.8	12.6	1603	14	BM808823	BM808823 AGENCOURT
	22	90.8	12.5	840	13	BI463235	BI463235 603203125
	23	90.6	12.4	1764	11	BC018330	BC018330 Mus muscu
	24	90.6	12.4	2611	11	AK004891	AK004891 Mus muscu
	25	90	12.3	1144	13	BM546496	BM546496 AGENCOURT
	26	89.2	12.2	634	12	BG714306	BG714306 602669812
	27	87	11.9	720	13	BI753905	BI753905 603022907
	28	86.8	11.9	637	10	BB642387	BB642387 BB642387
	29	86.8	11.9	906	13	BI873140	BI873140 603397789
	30	85.6	11.7	738	13	BI600170	BI600170 603246569
	31	84	11.5	612	12	BF446062	BF446062 7p18h12.x
	32	83.2	11.4	843	17	CNS04SGJ	AL305164 Tetraodon
	33	81.6	11.2	510	10	BE654347	BE654347 UI-M-AO0-
	34	81.6	11.2	843	17	AZ551618	AZ551618 ENTDV54TR
c	35	81.6	11.2	908	17	AZ548467	AZ548467 ENTEK30TR
c	36	81.6	11.2	931	17	BH160272	BH160272 ENTQV49TR
	37	80.4	11.0	633	10	BB642318	BB642318 BB642318
	38	80	11.0	768	12	BE746231	BE746231 601579712
	39	79.6	10.9	897	13	BI872266	BI872266 603396606
	40	79.4	10.9	663	10	BB656301	BB656301 BB656301
	41	79	10.8	1189	11	AK005368	AK005368 Mus muscu
	42	77.8	10.7	566	10	BE266564	BE266564 601193555
	43	77.6	10.6	2790	11	AK020259	AK020259 Mus muscu
	44	77.2	10.6	553	13	BJ499680	BJ499680 BJ499680
	45	77	10.6	741	13	BI224313	BI224313 602940621

ALIGNMENTS

RESULT 1
BM737726
LOCUS BM737726 540 bp mRNA linear EST 01-MAR-2002
DEFINITION K-EST0000769 S1SNU5 Homo sapiens cDNA clone S1SNU5-16-A04 5', mRNA sequence.
ACCESSION BM737726
VERSION BM737726.1 GI:19059055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 540)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 16 row: A column: 04
High quality sequence stop: 540.
FEATURES Location/Qualifiers
source 1. .540

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="S1SNU5-16-A04"
                     /clone_lib="S1SNU5"
                     /sex="F"
                     /tissue_type="Ascites"
                     /cell_type="Lymphoblast-like"
                     /cell_line="SNU-5"
                     /lab_host="Top10F'"
                     /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT       96 a    180 c    136 g    128 t
ORIGIN

  Query Match             72.4%;  Score 528;  DB 14;  Length 540;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-110;
  Matches  539;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3 GGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAGAA 62
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAGAA 60

Qy      63 ACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCACTT 122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCACTT 120

Qy     123 CTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGT 182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGT 180

Qy     183 CCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCT 242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCT 240

Qy     243 CTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTA 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTA 300

Qy     303 TGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGC 362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       301 TGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGC 360

Qy     363 CCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCG 422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCG 420

Qy     423 CTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCT 482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCT 480

Qy     483 CAGGATCCTCGGCATCGTCT-GGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCA 541
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     481 CAGGATCCTCGGCATCGTCTGGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCA 540


RESULT 2
AG046172/c
LOCUS       AG046172                 670 bp    DNA     linear   GSS 02-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-025E24.R, genomic survey sequence.
ACCESSION   AG046172
VERSION     AG046172.1  GI:16583064
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-025E24.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 670)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .670
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="PTB-025E24.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      161 a    157 c    213 g    138 t      1 others
ORIGIN

  Query Match             67.0%;  Score 488.4;  DB 17;  Length 670;
  Best Local Similarity   98.6%;  Pred. No. 2.1e-101;
  Matches  492;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     231 CAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCC 290
          |||||||||||||||||  |||||||||||||||||||| ||||||||||||||||||
Db     669 CAACTACTACCTCTTCAGCAGGGCGGTCTCTGACCTCCTGGTACTGCTCCTTGGAATGCC 610

Qy     291 CCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTA 350
          |||| ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     609 CCTGNAGGTCTATGAGATGTGGCGCAACTACCCTATCTTGTTCGGGCCCGTGGGCTGCTA 550

Qy     351 CTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGT 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     549 CTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGT 490

Qy     411 CAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCG 470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     489 CAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCG 430

Qy     471 GCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCC 530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     429 GCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCC 370

Qy     531 CAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGG 590
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     369 CAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCTTGGTCCCAGG 310

Qy     591 TTCGGCCACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCAC 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     309 TTCGGCCACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCAC 250
```

```
Qy     651 CTCCTTCCTATTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGC 710
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     249 CTCCTTCCTATTCTACCTCCTCCCAATGACTGTCATCAGTGTCCTCTACTACCTCATGGC 190

Qy     711 ACTCAGAGTGAGTATCTAG 729
           ||||||||||||||||||
Db     189 ACTCAGAGTGAGTATCTAG 171


RESULT 3
BE751626
LOCUS       BE751626                555 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  203809 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE751626
VERSION     BE751626.1  GI:10165618
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 44  row: O  column: 6
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      100 a    165 c    134 g    156 t
ORIGIN

  Query Match             48.6%;  Score 354.2;  DB 12;  Length 555;
  Best Local Similarity   87.0%;  Pred. No. 9.7e-71;
  Matches  389;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           |||||||||| | |||||| |||||||| ||||| ||| ||||  ||||||| ||| ||
Db     109 ATGGAAAAACATGAGAATGTTTCCTGGATGTACCAACAGGAACTGAAAGATCCTTTCAAG 168

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||| ||||||||| ||| || || || || || || || |||||||| ||||| |||||||||
Db     169 AAATACCTGAACAACACAGATGACTACCTAGCTTTGCTCTGCGGGCCTCGTCGCAGCCAC 228

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           |||||||||| ||| ||| |||||||| | |||||||||||||||||| ||||||||
Db     229 CTCTTCCTCCCGGTGACTGCGGTGTATGCTCTGATTTTTGTGGTGGGGGTCGTTGGCAAT 288

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           |||||||||| |||||||||| | ||||||| | |||||||| |||||||| ||||||
Db     289 CTCCTGGTGTGCTTGGTGATTCTTCGGCACCAGACGATGAAGACCCCCACCAATTACTAC 348

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||| |||| ||||| |||||||| ||| |||| || || |||||||||| |||
Db     349 CTCTTCAGCTTGGCTGTCTCCGACCTCCTAGTCTTGCTTCTCGGGATGCCCCTGGAAGTC 408

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||| ||| |||| ||||| ||||||||||||||||||||
Db     409 TATGAGATGTGGCGCAACTACCCCTTCCTGTTTGGGCCGGTGGGCTGCTACTTCAAGACG 468

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||| |||||||||| |||||||||||||||||| ||||||||||||||||||||
Db     469 GCCCTCTTCGAGACCGTGTGTTTCGCCTCCATCCTCAGCGTCACCACCGTCAGCGTGGAG 528

Qy     421 CGCTACGTGGCCATCCTACACCCGTTC 447
           ||||| || ||| |||| |||||||||
Db     529 CGCTATGTAGCCGTCCTCCACCCGTTC 555


RESULT 4
BE871983
LOCUS       BE871983                929 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601448032F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3852151 5',
            mRNA sequence.
ACCESSION   BE871983
VERSION     BE871983.1  GI:10320759
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9573  row: p  column: 08
            High quality sequence start: 11
            High quality sequence stop: 602.
FEATURES             Location/Qualifiers
     source          1. .929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3852151"
                     /clone_lib="NIH_MGC_65"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies.  "
BASE COUNT      186 a    252 c    228 g    263 t
ORIGIN

  Query Match             43.2%;  Score 315;  DB 12;  Length 929;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-61;
  Matches  481;  Conservative    0;  Mismatches   65;  Indels   17;  Gaps   11;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           |||||||||  ||| ||||||||||||||||||||||| |||||||| ||||||||||||
```

```
Db    143 ATGGAAAAATCTCAAAATGCTTCCTGGATCTACCAG-AGAAACTACAAGATCCATTCCAA 201

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCC-TTCCTCTGCGGACCTCGGCGCAGCCA 119
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    202 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTTCCTCTGCGGACCTCGGCGCAGCCA 261

Qy    120 CTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAA 179
          |||||||||||||||||||||||||||||||||| | |||| |||||||||| |||||
Db    262 CTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTATGGTGGGGGGGGTCATGGGCAA 321

Qy    180 TGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTA 239
          |||||||||||||||||| |||||||||||||||||||||||||||||| |||||
Db    322 TGTCCTGGTGTGCCTGGTGATCCTGCAGCACCAGGCTATGAAGACGCCCACCAA-TACTA 380

Qy    240 CCTCTTCAGCCTGG-CGGTCTCTGACCTCCTGG--TCCTGCTCCTTGGAATGCCCCTGGA 296
          |||||||||||||| |||||||||||||||||  || |||||||| | | | ||| |
Db    381 CCTCTTCAGCCTGGCCGGTCTCTGACCTCCTGGGTTCTTGCTCCTTTGGAAGGCCCCGTG 440

Qy    297 GGTCT---ATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCC---GTGGGCTGCTA 350
          |   |   |||| |||||||||||||||| || |||    || ||     ||||||
Db    441 GAGGTCTAATGACATGTGGCGCAACTACCCCTTTTTGGGTCGGGCCCGTTGGGGCTGTAC 500

Qy    351 CTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGT 410
          |||||| |||||| |||||||||||||||||||||||||||||||| |||| ||||||
Db    501 ATTCAAGAGGGCCCTTTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGAATCA-CACCGT 559

Qy    411 CAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCG 470
          ||| |||     |||||||| |||||||||||||| ||||||||| ||  |||||||||
Db    560 CAG-GTGAGCGCCTACGTGGCAATCCTACACCCGTT-CGCGCCAAATTGAGGAGCACCCG 617

Qy    471 GCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCC 530
          |   | || ||||||||  ||||| ||  |||||||| ||||| || ||||||||||
Db    618 GGTGCCGGGCCTCAGGAATCTCGGAATTCTCTGGGGCTACTCCGAGCACTTCTCCCTG-- 675

Qy    531 CAACACCAGCATCCATGGCATCA 553
          ||| |||||||||   ||| |||
Db    676 CAAAACCAGCATCATGGGCTTCA 698
```

RESULT 5
BB587450
LOCUS BB587450 624 bp mRNA linear EST 26-OCT-2001
DEFINITION BB587450 RIKEN full-length enriched, adult female vagina Mus musculus cDNA clone 9930001G13 5', mRNA sequence.
ACCESSION BB587450
VERSION BB587450.2 GI:16449720
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 624)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT On Nov 30, 2000 this sequence version replaced gi:11483994.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES Location/Qualifiers
source 1. .624
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="9930001G13"
/clone_lib="RIKEN full-length enriched, adult female vagina"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 121 a 166 c 165 g 172 t
ORIGIN

Query Match 38.4%; Score 280; DB 10; Length 624;
Best Local Similarity 77.1%; Pred. No. 9.4e-54;
Matches 373; Conservative 0; Mismatches 95; Indels 16; Gaps 2;

```
Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          |||| ||||||| | |||||||||||||| ||               ||| |  ||  |
Db    106 ATGGGAAAACTTGAAAATGCTTCCTGGATCCAC--------------GATTCTCTCATG 150

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCT-TCCTCTGCGGACCTCGGCGCAGCCA 119
          ||  || |||||||||| ||||||||  ||||||  ||| || |||||   ||||||  |
Db    151 AAGTACTTGAACAGCACAGAGGAGTACTTGGCCTACCCTGTGTGGACCCAAGCGCAGTGA 210

Qy    120 CTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAA 179
          | | | ||| || ||||||||||| |||| ||  || || |||||||||| || |||||
Db    211 CCTATCCCTTCCAGTGTCTGTGGTCTATGCGCTGATCTTCGTGGTGGGGGTGATAGGCAA 270
```

```
Qy    180 TGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTA 239
          | | ||||||||||||||||||    |  || ||  || ||||||| ||||||||||||||
Db    271 TCTTCTGGTGTGCCTGGTGATTGCCCGACATCAAACTTTGAAGACACCCACCAACTACTA 330

Qy    240 CCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGT 299
          ||||||||| |||| |||| ||| | ||||||||||| | || |||| ||||||||
Db    331 TCTCTTCAGCTTGGCAGTCTCAGACTTGCTGGTCCTGCTCTTAGGTATGCCACTGGAGGT 390

Qy    300 CTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGAC 359
          ||| ||| |||||| ||| || || ||| |||| ||||| ||||| ||||||||||||||
Db    391 CTACGAGTTGTGGCACAATTATCCCTTCCTGTTTGGGCCGGTGGGATGCTACTTCAAGAC 450

Qy    360 GGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGA 419
          ||||| || ||||| |||||||| |||||||| |||||  ||||||| || ||| | ||
      451 AGCCCTTTTCGAGACTGTGTGCTTTGCCTCCATTCTCAGTGTCACCACGGTTAGCATTGA 510

Qy    420 GCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGC 479
          |||||||||||||||  | || |||||||| ||||| ||| ||||||| |||      ||
Db    511 GCGCTACGTGGCCATTGTCCATCCGTTCCGAGCCAAGCTGGAGAGCACACGGGCGACGGG 570

Qy    480 CCTC 483
          || |
Db    571 CCCC 574


RESULT 6
BE751310
LOCUS       BE751310                 465 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  203132 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE751310
VERSION     BE751310.1  GI:10165239
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 43  row: O  column: 6
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .465
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT       88 a    129 c    113 g    135 t
ORIGIN

  Query Match             37.8%;  Score 275.4;  DB 12;  Length 465;
  Best Local Similarity   85.7%;  Pred. No. 9.5e-53;
  Matches  306;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          |||||||||| | |||||| |||||||| ||||| ||| ||||  ||||||| ||| ||
Db    109 ATGGAAAAACATGAGAATGTTTCCTGGATGTACCAACAGGAACTGAAAGATCCTTTCAAG 168

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||| ||||||||| ||| || || || || || || |||||||| ||||| |||||||||
Db    169 AAATACCTGAACAACACAGATGACTACCTAGCTTTGCTCTGCGGGCCTCGTCGCAGCCAC 228

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          |||||||||| ||| ||| |||||||| |  ||||||||||||||||| |||||||||
Db    229 CTCTTCCTCCCGGTGACTGCGGTGTATGCTCTGATTTTTGTGGTGGGGGTCGTTGGCAAT 288

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          |||||||||| |||||||||| | ||||||| | |||||||| |||||||| ||||||
Db    289 CTCCTGGTGTGCTTGGTGATTCTTCGGCACCAGACGATGAAGACCCCCACCAATTACTAC 348

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||| |||| ||||| |||||||| ||| |||| || || |||||||||| |||
Db    349 CTCTTCAGCTTGGCTGTCTCCGACCTCCTAGTCTTGCTTCTCGGGATGCCCCTGGAAGTC 408

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAG 357
          ||||||||||||||||||||| ||| |||| ||||| ||||||||||||||||||
Db    409 TATGAGATGTGGCGCAACTACCCCTTCCTGTTTGGGCCGGTGGGCTGCTACTTCAAG 465


RESULT 7
BQ721315
LOCUS       BQ721315                 993 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8296894 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6194542 5', mRNA sequence.
ACCESSION   BQ721315
VERSION     BQ721315.1  GI:21860212
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 993)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13599  row: o  column: 23
            High quality sequence stop: 584.
FEATURES             Location/Qualifiers
     source          1. .993
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6194542"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
```

5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

BASE COUNT 160 a 326 c 308 g 198 t 1 others
ORIGIN

Query Match 31.6%; Score 230.2; DB 14; Length 993;
Best Local Similarity 66.9%; Pred. No. 2.8e-42;
Matches 342; Conservative 0; Mismatches 168; Indels 1; Gaps 1;

```
Qy    103 GGACCTCGGCGCAGCCACTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTG 162
          || || | ||  |   |  | ||| | ||| | | ||     ||  |||  || || |||
Db    220 GGGCCCCAGCAGACAGAGCTGTTCATGCCCATCTGTGCCACATACCTGCTGATCTTCGTG 279

Qy    163 GTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAG 222
          ||||| |   | |||||||  |||   || |||| || ||||  ||| |||| |||
Db    280 GTGGGCGCTGTGGGCAATGGGCTGACCTGTCTGGTCATCCTGCGCCACAAGGCCATGCGC 339

Qy    223 ACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTT 282
          ||||| |||||||||||||||||||||||||||| || || ||||| ||||| |||||  |
Db    340 ACGCCTACCAACTACTACCTCTTCAGCCTGGCCGTGTCGGACCTGCTGGTGCTGCTGGTG 399

 y    283 GGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTG 342
          ||  |||||||||||| |||||||||||||| |||||||| ||| || | ||    |
Db    400 GGCCTGCCCCTGGAGCTCTATGAGATGTGGCACAACTACCCCTTCCTGCTGGGCGTTGGT 459

Qy    343 GGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATC 402
          |||||||| |||   |||   || |||||||  || ||| | |||||  | |||| | ||
Db    460 GGCTGCTATTTCCGCACGCTACTGTTTGAGATGGTCTGCCTGGCCTCAGTGCTCAACGTC 519

Qy    403 ACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCA-AACTGCA 461
          ||  || | |||||||| ||||| |||||| |  | |||||  |||  ||||   |
Db    520 ACTGCCCTGAGCGTGGAACGCTATGTGGCCGTGGTGCACCCACTCCAGGCCAGGTCCATG 579

Qy    462 GAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTT 521
          |   || |||  ||  |  | | |  | || ||   |||||||||  |  || ||||||
Db    580 GGTGACGCGGGCCCATGTGCGCCGAGTGCTTGGGGCCGTCTGGGGTCTTGCCATGCTCTG 639

Qy    522 CTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCT 581
          |||||||||||||||||| | || ||||||  |   |    | ||||    ||  ||
Db    640 CTCCCTGCCCAACACCAGCCTGCACGGCATCCGGCAGCTGCCGTGCCCCTGCCGGGGCCC 699

Qy    582 GGTCCCAGGTTCGGCCACCTGTACGGTCATC 612
           || ||||  || ||    || | | |  ||
Db    700 AGTGCCAGACTCAGCTGTTTGCATGCTGGTC 730
```

RESULT 8
CNS036WH/c
LOCUS CNS036WH 941 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 216F22 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL230570
VERSION AL230570.1 GI:7889565
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 941)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 941)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 941)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .941
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="216F22"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG216DC11LP1~end : T7"
BASE COUNT 197 a 277 c 327 g 132 t 8 others
ORIGIN

Query Match 29.0%; Score 211.2; DB 17; Length 941;
Best Local Similarity 67.3%; Pred. No. 6e-38;
Matches 325; Conservative 1; Mismatches 155; Indels 2; Gaps 2;

```
Qy     73 AGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCACTTCTTCCTCCCC 132
          |||   |||| ||| ||||||     | | || || |   |    | | ||||||| |||
Db    568 AGCGAGGAGGCGTACCTGGCCAGGTACCGGGGCCCCCCCCAGGTCGCCCGTCTTCCTGCCC 509

Qy    133 GTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGC 192
          || |  || |  ||  || | || ||  |||||||||||| | |||||   ||||   |||
Db    508 GTCTGCGTCGCCTACCTGGCCATCTTCCTGGTGGGGGTCCTGGGCAACTCCCTGACCTGC 449

Qy    193 CTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTG 252
             || || || |    || ||| | |   |||||||||||||||||||  | | ||||
Db    448 GCCGTCATCCTCCGCTTCCGGGCGANGCGCACGCCCACCAACTACTACCTGCTGA-CCTG 390

Qy    253 GCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGG 312
          ||||  || ||||| ||||| ||||| || || ||||| |||||| | || ||  |||||
Db    389 GCGGCGTCCGACCTGCTGGTGCTGCTGCTGGGGATGCCGCTGGAGCTGTACGACCTGTGG 330

Qy    313 CGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAG 372
          || |||||||| ||| || | |||   | ||||||||||||  |||   ||| || |||
Db    329 CGGAACTACCCCTTCCTGCTGGGGGAGGGGGGCTGCTACTTCCGGACCTTCCTGTTCGAG 270

Qy    373 ACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCC 432
          || |||| :|||||||||||||||| | ||||  |  |||||||||||||||| |||||
Db    269 ACGGTGT-STTCGCCTCCATCCTCAACGTCACGGCGCTCAGCGTGGAGCGCTACCTGGCC 211

Qy    433 ATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTC 492
            |  | ||||||  ||   | ||| |    |   || ||    |  | |  | | |  ||
Db    210 GTGGTGCACCCGCTCAAGGTCAAGCACATGGCAACTCGCGCGCACGTCAAGAAGGTGGTC 151

Qy    493 GGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATC 552
            | |  | |||    | ||| |||| | | |  |||||||||||||| | || ||||||
Db    150 CTCCTGCTGTGGCTGCTGTCCATGCTGTGCGCGGTGCCCAACACCAGCGTGCACGGCATC 91

Qy    553 AAG 555
           ||
Db     90 GAG 88
```

RESULT 9
AU244292
LOCUS AU244292 624 bp mRNA linear EST 15-JAN-2002
DEFINITION AU244292 Medaka eye cDNA library (SNK01) Oryzias latipes cDNA clone NGY47.02e similar to pir|T15816|hypothetical protein C48C5.1 - Caenorhabditis elegans, mRNA sequence.
ACCESSION AU244292
VERSION AU244292.1 GI:18156871

```
KEYWORDS    EST.
SOURCE      Japanese medaka.
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Sanaka,E., Hori,H., Naruse,K., Mitani,H. and Tanaka,M.
  TITLE     Medaka EST analysis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Emi Sanaka
            Department of Biological Sciences
            Graduate School of Science, Nagoya University
            Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
            Tel: 81-52-789-2973
            Fax: 81-52-789-2974
            Email: sanaka@bio.nagoya-u.ac.jp
            This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Oryzias latipes"
                     /strain="wild type"
                     /db_xref="taxon:8090"
                     /clone="NGY47.02e"
                     /clone_lib="Medaka eye cDNA library (SNK01)"
                     /tissue_type="eye"
                     /dev_stage="adult"
                     /note="Wild samples from Okayama Pref.(Southern part of
                     Japan)"
BASE COUNT      118 a    194 c    148 g    164 t
ORIGIN

  Query Match             27.5%;  Score 200.6;  DB 9;  Length 624;
  Best Local Similarity   66.1%;  Pred. No. 1.4e-35;
  Matches  290;  Conservative    0;  Mismatches  149;  Indels    0;  Gaps    0;

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||| |   ||| ||  | |  || ||| || | ||  |   ||| |||
Db    178 TTCTTCCTCCCCGTGACGACGGTTTACCTTCTCATCTTTCTGATTGGCCTGTCTGGAAAT 237

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          || ||    ||   ||||||     | ||| ||   |||| ||  |||||||||| ||||
 b    238 GTGCTCACCTGTGCGGTGATAGCAAAACACAAGAGGATGAGGAACCCCACCAACTTCTAC 297

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||  | ||||||||  || ||  |||||||| || ||| |  | || ||||||  ||||| |
Db    298 CTTGTGAGCCTGGCCGTGTCGGACCTCCTAGTGCTGATGTTCGGGATGCCCTTGGAGATT 357

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          || ||  ||||||  |||||||| |||   |||||    |  ||||||||||||||||
Db    358 TACGACCTGTGGCAGAACTACCCGTTCCCCTTCGGCGAGGGCGGCTGCTACTTCAAGACC 417

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
            |||||| ||||| || ||||||||||||||||||| | ||||  |  | |||||||||
Db    418 TTCCTCTTCGAGACGGTCTGCTTCGCCTCCATCCTCAACGTCACAGCGCTGAGCGTGGAG 477

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          || ||| | ||| |   | || ||  ||    ||      | |  |||| |  |  |||
Db    478 CGGTACATCGCCGTGGTGCATCCTCTCAAAACCCGTTACCTGTCCACCAACCAGCATGCC 537

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
            |  || ||    | || ||||   | ||  || |||   || | |||||||||  |
Db    538 AAGCGCGTCATCACAGTGGTGTGGGTGGTGTCAATGGTCTGTGCCATCCCCAACACCTCC 597

Qy    541 ATCCATGGCATCAAGTTCC 559
          | |||||||||   | ||
Db    598 CTGCATGGCATCTTCTACC 616


RESULT 10
AU244285
LOCUS       AU244285                445 bp    mRNA    linear   EST 15-JAN-2002
DEFINITION  AU244285 Medaka eye cDNA library (SNK01) Oryzias latipes cDNA clone
            NGY47.01f, mRNA sequence.
ACCESSION   AU244285
VERSION     AU244285.1  GI:18156864
KEYWORDS    EST.
SOURCE      Japanese medaka.
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   Sanaka,E., Hori,H., Naruse,K., Mitani,H. and Tanaka,M.
  TITLE     Medaka EST analysis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Emi Sanaka
            Department of Biological Sciences
            Graduate School of Science, Nagoya University
            Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
            Tel: 81-52-789-2973
            Fax: 81-52-789-2974
            Email: sanaka@bio.nagoya-u.ac.jp
            This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.
FEATURES             Location/Qualifiers
     source          1. .445
                     /organism="Oryzias latipes"
                     /strain="wild type"
                     /db_xref="taxon:8090"
                     /clone="NGY47.01f"
                     /clone_lib="Medaka eye cDNA library (SNK01)"
                     /tissue_type="eye"
                     /dev_stage="adult"
                     /note="Wild samples from Okayama Pref.(Southern part of
                     Japan)"
BASE COUNT       85 a    132 c    104 g    123 t      1 others
ORIGIN

  Query Match             22.9%;  Score 166.6;  DB 9;  Length 445;
  Best Local Similarity   71.0%;  Pred. No. 7e-28;
  Matches  220;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          |||||||||||||||  |  ||| ||  | |  || ||| || | ||  |   ||| |||
Db    136 TTCTTCCTCCCCGTGNTTACGGTTTACCTTCTCATCTTTCTGATTGGCCTGTCTGGAAAT 195

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          || ||    ||   ||||||     | ||| ||   |||| ||  |||||||||| ||||
Db    196 GTGCTCACCTGTGCGGTGATAGCAAAACACAAGAGGATGAGGAACCCCACCAACTTCTAC 255

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||  | ||||||||  || ||  |||||||| || ||| |  | || ||||||  ||||| |
Db    256 CTTGTGAGCCTGGCCGTGTCGGACCTCCTAGTGCTGATGTTCGGGATGCCCTTGGAGATT 315

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          || ||  ||||||  |||||||| |||   |||||    |  ||||||||||||||||
Db    316 TACGACCTGTGGCAGAACTACCCGTTCCCCTTCGGCGAGGGCGGCTGCTACTTCAAGACC 375

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
            |||||| ||||| || ||||||||||| |||||||| | ||||  |  | |||||||||
Db    376 TTCCTCTTCGAGACGGTCTGCTTCGCCTTCATCCTCAACGTCACAGCGCTGAGCGTGGAG 435

Qy    421 CGCTACGTGG 430
          || ||| | |
Db    436 CGGTACATCG 445


RESULT 11
CNS04GXP
LOCUS       CNS04GXP                959 bp    DNA     linear   GSS 21-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            109K06 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL290230
```

```
VERSION     AL290230.1  GI:8028810
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 959)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 959)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="109K06"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG109BF03SP1~end :
                     PUC-Ori"
BASE COUNT      184 a    293 c    290 g    184 t      8 others
ORIGIN

  Query Match             16.0%;  Score 116.4;  DB 17;  Length 959;
  Best Local Similarity   55.9%;  Pred. No. 2.9e-16;
  Matches  254;  Conservative    4;  Mismatches  190;  Indels    6;  Gaps    2;

Qy    120 CTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAA 179
          |||| |    | | | |      ||    |||: |      | ||||||  |   |  |
Db    378 CTTCCTGACGCTCCTCTACAGCCTGATGCTGCYGACGGGAATCGTGGGGAACGCGGTGAC 437

Qy    180 TGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTA 239
            ||  |||| ||| |||| |   |||| ||       || ||| |  |  || | || |
Db    438 CATCAGGGTGAGCCAGGTGCTGAAGCAGAACGGCTACCTGCAGAAGAACGTCACCGACCA 497

      240 CCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGT 299
          | |  ||||||||||   ||| ||||||||||| ||||||  | || |||||  ||||| |
Db    498 CATGATCAGCCTGGCCTGCTCCGACCTCCTGGTGCTGCTCATCGGCATGCCGGTGGAGCT 557

Qy    300 CTA---TGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAA 356
          |||    |  || |||  |  || | |    || |||  |||  |||  |  :| |
Db    558 CTACAGCGCCATCTGGTTCCCCTTCACCTCGGCGTCCGGCAACGTCTCCTGTAAGAWCTA 617

Qy    357 GACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGT 416
          |    ||| || ||| |     ||| |||| ||||  | | |||  |||| | ||| |
Db    618 CAACTTCCTGTTCGAGGCGTGCAGCTACGCCACCATTGTGAACATAGCCACGCTTAGCTT 677

Qy    417 GGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCG 476
          ||||||:|||||||||||   ||||| ||||||  |||    | | |   | ||   ||
Db    678 CGAGCGCYACGTGGCCATCTGCCACCCCTTCCGCTACAAGGCCCTGGG---CGGGAAGCG 734

Qy    477 GGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACAC 536
            ||  |    || ||| | ||| ||||    | |: |||||  |  |||||||
Db    735 CACCAACGCCCTCATCGCCTTCGCCTGGCTGGTGTYGGTGCTGGTGGCCCTGCCGCYGCT 794

Qy    537 CAGCATCCATGGCATCAAGTTCCACTACTTCCCC 570
          || |  |   || |   ||   |||  |  || |
Db    795 CATCGCCACGGGAACGCAGGGGCACATCCCCCTC 828


RESULT 12
CNS01XJ2/c
LOCUS       CNS01XJ2                817 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            202F23 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL171767
VERSION     AL171767.1  GI:7809824
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 817)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 817)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 817)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .817
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="202F23"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG202CC12SP1~end :
                     PUC-Ori"
BASE COUNT      158 a    237 c    258 g    159 t      5 others
ORIGIN

  Query Match             14.8%;  Score 107.8;  DB 17;  Length 817;
  Best Local Similarity   58.6%;  Pred. No. 2.5e-14;
  Matches  242;  Conservative    1;  Mismatches  163;  Indels    7;  Gaps    3;

Qy    161 TGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGA 220
          | ||||||  |   |  |   ||  |||| ||| |||| |   |||| ||       ||
Db    481 TCGTGGGGAACGCGGTGACCATCAGGGTGAGCCAGGTGCTGAAGCAGAACGGCTACCTGC 422

Qy    221 AGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCC 280
          ||| |  |  || | || || |  ||||||||||   ||| ||||||||||| ||||||
Db    421 AGAAGAACGTCACCGACCACATGATCAGCCTGGCCTGCTCCGACCTCCTGGTGCTGCTCA 362

Qy    281 TTGGAATGCCCCTGGAGGTCTA---TGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGC 337
          | || |||||  |||||  ||||     |  || |||  |  || | | |    || |||
Db    361 TCGGCATGCCGGTGGAGCTCTACAGCGCCATCTGGTTCCCCTTCACCTCGGCGTCCGGCA 302

Qy    338 CCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCA 397
           |:|   |||  |  || | |    ||| || |||  ||||   | |||| ||||  | |
Db    301 ACKTCTCCTGTAAGATCTACAACTTCCTGTTCGAG-GCGTGCAGTACGCCACCATTGTGA 243
```

```
Qy     398 GCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAAC 457
           |||  ||||  | ||| | ||||||||||||||||||   ||||| ||||||  |||
Db     242 ACATAGCCACGCTTAGCTTCGAGCGCTACGTGGCCATCTGCCACCCCTTCCGCTACAAGG 183

Qy     458 TGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGC 517
           | | |   | ||   ||  ||  |    || ||| | ||| ||||    | || ||||
Db     182 CCCTGGG---CGGGAAGCGCACCAACGCCCTCATCGCCTTCGCCTGGCTGGTGTCGGTGC 126

Qy     518 TCTTCTCCCTGCCCAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCC 570
           |  |  |||||||      || |  |   || |   ||   |||  |  || |
Db     125 TGGTGGCCCTGCCGCTGCTCATCGCCACGGGAACGCAGGGGCACATCCCCCTC 73


RESULT 13
BF603623
LOCUS       BF603623                500 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  269181 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BF603623
VERSION     BF603623.1  GI:11701421
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 49  row: I  column: 6
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .500
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT       71 a    179 c    139 g    111 t
ORIGIN

  Query Match             14.7%;  Score 107;  DB 12;  Length 500;
  Best Local Similarity   55.8%;  Pred. No. 3.1e-14;
  Matches  225;  Conservative    0;  Mismatches  175;  Indels    3;  Gaps    1;

Qy     132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
           |||  | |     | ||| |  | || ||||||||  ||   |||||  | ||
Db      67 CGTCACAGCCACCTGCGTGGCGCTCTTCGTGGTGGGCATCGCGGGCAACCTGCTCACCAT 126

Qy     192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
           |||||| |   ||    ||  |   ||   ||  ||||||||  |||||| | |||| |

Db     127 GCTGGTGGTGTCGCGTTTCCGCGAGCTGCGTACCACCACCAACCTCTACCTGTCCAGCAT 186

Qy     252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
           |||  |||| ||| | |   || |  |||  | |||||||| ||  ||       | ||
Db     187 GGCCTTCTCCGACTTAC---TCATCTTCCTCTGCATGCCCCTCGACCTCGTGCGCCTCTG 243

Qy     312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
           ||   ||   || |     || ||   | ||  |||| |  ||     |    ||  ||
Db     244 GCATTACCGGCCCTGGAACTTGGGCGACCTGCTCTGCAAACTCTTCCAGTTTGTCAGCGA 303

Qy     372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
           || |      || |||  |  | |||| |||||||  |  | ||||| |||||||||  | ||
Db     304 GAGCTGCACCTACGCTACGGTGCTCACCATCACCGCGCTGAGCGTCGAGCGCTACTTCGC 363

Qy     432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
           ||||      |||| | || |||||  ||  || ||||  | ||||||          || |
Db     364 CATCTGCTTCCCGCTGCGGGCCAAGGTGGTGATCACCAAGGGCCGGGTGAAGCTCGTCAT 423

Qy     492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAAC 534
           |    || |||||| | |  |  |   |  | ||  |||||
Db     424 CCTGGTCATCTGGGCCGTGGCTTTCTGCAGCGCCGGGCCCATC 466


RESULT 14
BM543497
LOCUS       BM543497                1027 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6492823 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726591
            5', mRNA sequence.
ACCESSION   BM543497
VERSION     BM543497.1  GI:18773939
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1027)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12719  row: e  column: 24
            High quality sequence stop: 635.
FEATURES             Location/Qualifiers
     source          1. .1027
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5726591"
                     /clone_lib="NIH_MGC_124"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
                     (destroyed); Site_2: NotI; RNA source male hippocampus,
                     age 27. Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.4 kb, insert size range 0.9-4 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 012."
BASE COUNT      153 a    357 c    337 g    180 t
ORIGIN

  Query Match             13.4%;  Score 97.6;  DB 13;  Length 1027;
  Best Local Similarity   51.9%;  Pred. No. 5.9e-12;
  Matches  295;  Conservative    0;  Mismatches  264;  Indels    9;  Gaps    3;
```

```
Qy    137 CTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGG 196
          | | | | || |  |  || |  | | |||| |     |||||||  |||      |  |
Db    150 CCGCGCTCTACGCACTCATCTGGGCGCTGGGCGCGGCGGGCAATGCGCTGTCCGTGCACG 209

Qy    197 TGATTCTGCAGCACCAGGCTATG---AAGACGCCCACCAACTACTACCTCTTCAGCCTGG 253
          || | ||| ||   | |||   |     |  || |     | || || |  |||||||||
Db    210 TGGTGCTGAAGGCGCGGGCCGGGCGCGCGGGGCGCCTGCGCCACCACGTGCTCAGCCTGG 269

Qy    254 CGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGA---TGT 310
          || || | | ||| ||| | |||||  | ||  ||||  |||||  ||||       |||
Db    270 CGCTCGCGGGCCTGCTGCTGCTGCTGGTCGGCGTGCCGGTGGAGCTCTACAGCTTCGTGT 329

Qy    311 GGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTG 370
          ||  | |||||||| |   | |||||   | ||||||||  |  | |       | |    |
Db    330 GGTTCCACTACCCCTGGGTCTTCGGCGACCTGGGCTGCCGCGGCTACTACTTCGTGCACG 389

Qy    371 AGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGG 430
          ||        || |||| |  | || ||| |  |   | | ||||  ||||||| | | |
Db    390 AGCTGTGCGCCTACGCCACGGTGCTGAGCGTGGCAGGCCTGAGCGCCGAGCGCTGCCTAG 449

Qy    431 CCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCC 490
          || |    || ||  | || |||       | |   || |  |||||| |||   || |
Db    450 CCGTGTGCCAGCCCCTGCGTGCCCGCAGCCTGCTGACGCCACGCCGGACCCGGTGGCTGG 509

Qy    491 TCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCA 550
          | |   ||   |||| |  |||  |   | || |||||||||   ||  ||||   ||
Db    510 TGGCGCTCTCGTGGGCCGCCTCGCTCGGCCTCGCCCTGCCCATGGCCGTCATCATGGGGC 569

Qy    551 TCAAGTTCCA---CTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGG 607
            |||  | |   | |     |  | |||   | ||  || |  |||      || ||||
Db    570 AGAAGCACGAACTCGAGACGGCGGACGGGGAGCCGGAGCCCGCCTCGCGAGTGTGCACGG 629

Qy    608 TCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACC 667
          |  |   | |           | |  || | |||||||| |     |  | | || |
Db    630 TGCTGGTGAGCCGCACCGCGCTCCAAGTCTTTATCCAGGTGAATGTGCTGGTGTCCTTCG 689

Qy    668 TCCTCCCCATGACTGTCATCAGTGTCCT 695
          | |||||| || |  | |    | ||||
Db    690 TGCTCCCCTTGGCACTAACTGCTTTCCT 717
```

RESULT 15
BF323227
LOCUS BF323227 448 bp mRNA linear EST 21-NOV-2000
DEFINITION maa38a07.y1 NCI_CGAP_Brn63 Mus musculus cDNA clone IMAGE:3813037 5' similar to SW:NTR2_RAT Q63384 NEUROTENSIN RECEPTOR TYPE 2 ;, mRNA sequence.
ACCESSION BF323227
VERSION BF323227.1 GI:11272778
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 448)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Other_ESTs: maa38a07.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

MGI:1455149
Seq primer: -40RP from Gibco

High quality sequence stop: 416.

FEATURES Location/Qualifiers
source 1. .448
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3813037"
/clone_lib="NCI_CGAP_Brn63"
/sex="female"
/dev_stage="10 weeks"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 53 a 157 c 148 g 89 t 1 others
ORIGIN

Query Match 13.1%; Score 95.4; DB 12; Length 448;
Best Local Similarity 55.2%; Pred. No. 1.3e-11;
Matches 208; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

```
Qy    187 GTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTC 246
          |||  | |||||  |   |  ||    || |        || |     |||| || |  ||
Db      8 GTGCACGTGGTGCTGAAGGCGCGGGCGGGTCGCCCCGGGCGCCTGCGCTACCACGTGCTC 67

Qy    247 AGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAG 306
          ||||||||  | || | ||| ||| | |||||  |  |  ||||| ||||| ||||  |
Db     68 AGCCTGGCACTGTCAGCCCTGCTGCTACTGCTGATCAGCGTGCCCATGGAGCTCTACAAC 127

Qy    307 A---TGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCC 363
            |||||  | |||||||| |   | ||| |     | |||||      | |      |
Db    128 TTCGTGTGGTCCCACTACCCCTGGGTCTTCTGCGATCTCGGCTGTCGTGGCTATTACTTC 187

Qy    364 CTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGC 423
            |   |||         || |||| |  | || ||| |  ||| | | |||||  ||||||
Db    188 GTGCGCGAGCTGTGCGCCTACGCCACGGTGCTGAGCGTGGCCAGCCTGAGCGCAGAGCGC 247

Qy    424 TACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTC 483
          | | ||||| |    || ||| | ||||||   |  | |  ||||| ||||||  ||| |
Db    248 TGCCTGGCCGTGTGCCAGCCGCTGCGCGCCCGCCGCCTGCTCACCCCGCGCCGCACCCGC 307

Qy    484 AGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATC 543
            |  | ||     | |||||||  | |||  ||  | |  |||||||||   |    |||
Db    308 CGCCTGCTGTCACTGGTCTGGGTCGCCTCTCTGGGCCTTGCCCTGCCCATGGCGGTTATC 367

Qy    544 CATGGCATCAAGTTCCA 560
            ||    |||  | |
Db    368 ATGGGACAGAAGCACGA 384
```

Search completed: January 17, 2003, 05:02:49
Job time : 2233 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:28:10 ; Search time 59 Seconds
(without alignments)
5509.118 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc.....cactcagagtgagtatctag 729

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119.8	16.4	1258	10	US-09-804-551B-25	Sequence 25, Appl
2	93	12.8	1427	10	US-09-967-768A-296	Sequence 296, App
3	89.4	12.3	1362	9	US-09-970-966-208	Sequence 208, App
4	89.4	12.3	1362	10	US-09-825-294-208	Sequence 208, App
5	85.6	11.7	1647	10	US-09-829-631A-12	Sequence 12, Appl
6	82.6	11.3	2108	10	US-09-917-800A-1620	Sequence 1620, Ap
7	82	11.2	1344	9	US-10-077-870-1	Sequence 1, Appli
8	82	11.2	1344	10	US-09-825-923-1	Sequence 1, Appli
9	82	11.2	1353	9	US-10-077-870-3	Sequence 3, Appli
10	82	11.2	1353	10	US-09-825-923-3	Sequence 3, Appli
11	79	10.8	2487	12	US-10-044-090-444	Sequence 444, App
12	77.8	10.7	2048	12	US-10-052-589-1	Sequence 1, Appli
13	77.8	10.7	2595	10	US-09-964-824A-249	Sequence 249, App
14	77	10.6	2050	9	US-09-940-240-15	Sequence 15, Appl
15	77	10.6	2625	9	US-09-954-531-995	Sequence 995, App
16	76.8	10.5	1080	10	US-09-739-151-1	Sequence 1, Appli
17	76.8	10.5	1679	9	US-09-104-063-5	Sequence 5, Appli
18	75.2	10.3	1047	9	US-10-090-569-6	Sequence 6, Appli
19	75.2	10.3	1805	10	US-09-823-114-18	Sequence 18, Appl
20	75.2	10.3	1829	9	US-09-905-186A-9	Sequence 9, Appli
21	75.2	10.3	1829	9	US-09-905-186A-10	Sequence 10, Appl
22	75.2	10.3	1829	9	US-09-905-186A-11	Sequence 11, Appl
23	75.2	10.3	1882	9	US-10-090-569-11	Sequence 11, Appl
24	74.8	10.3	2080	12	US-10-044-090-628	Sequence 628, App
25	74.8	10.3	2081	12	US-10-044-090-627	Sequence 627, App
26	74.8	10.3	2602	9	US-09-905-186A-1	Sequence 1, Appli
27	74.4	10.2	1914	10	US-09-829-631A-7	Sequence 7, Appli
28	74.4	10.2	2108	10	US-09-829-631A-9	Sequence 9, Appli
29	73.8	10.1	1670	10	US-09-880-107-3833	Sequence 3833, Ap
30	73.8	10.1	1895	9	US-09-940-240-19	Sequence 19, Appl
31	73.6	10.1	1829	9	US-09-905-186A-7	Sequence 7, Appli
32	73.6	10.1	1829	9	US-09-905-186A-8	Sequence 8, Appli
c 33	73.4	10.1	720	9	US-10-101-487-74	Sequence 74, Appl
34	73.4	10.1	720	9	US-10-101-487-76	Sequence 76, Appl
35	70.6	9.7	1773	9	US-10-112-599A-3	Sequence 3, Appli
36	70.4	9.7	1296	10	US-09-826-508-21	Sequence 21, Appl
37	70.4	9.7	1296	10	US-09-899-532-1	Sequence 1, Appli
38	70.4	9.7	1710	10	US-09-899-532-3	Sequence 3, Appli
39	69.6	9.5	823	12	US-10-044-090-598	Sequence 598, App
40	68.4	9.4	1197	10	US-09-796-338A-18	Sequence 18, Appl
41	68.4	9.4	1197	10	US-09-897-201-1	Sequence 1, Appli
42	68.4	9.4	1609	10	US-09-796-338A-16	Sequence 16, Appl
43	67.2	9.2	1272	8	US-08-796-570A-2	Sequence 2, Appli
44	67.2	9.2	1272	9	US-10-067-477-2	Sequence 2, Appli
45	66.4	9.1	1149	10	US-09-993-844-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-804-551B-25
; Sequence 25, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804,551B
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 1258
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1257)
US-09-804-551B-25

Query Match 16.4%; Score 119.8; DB 10; Length 1258;
Best Local Similarity 54.4%; Pred. No. 1.9e-21;
Matches 241; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

```
Qy    120 CTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAA 179
          | || |  | |||||  | ||||| ||    |  |||||  |    || ||  | |||||
Db     54 CATCGTGATACCCGTAACGGTGGTCTACTCCCTCATTTTCATAACCGGTGTAGTGGGCAA 113

Qy    180 TGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTA 239
            |      ||| | ||||||  | || |||   | ||| | ||| |||| || |||||
Db    114 CATAAGTACCTGCATTGTGATTAAGAAGAACCGTTCAATGCACACGGCCACGAATTACTA 173

Qy    240 CCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGT 299
          ||||||    |||||  |||| ||  ||||| ||||| |    ||  | || | ||||||
Db    174 CCTCTTTTCGCTGGCCATCTCGGATTTCCTGCTCCTGTTGTCGGGCGTTCCGCAGGAGGT 233

Qy    300 CTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGAC 359
            |   | || |||  ||| ||||| | | |||| |||        ||||  |
Db    234 GTCCTACATCTGGTCCAAGTACCCGTACGTGTTTGGGGAGTACATCTGCATCGGACGTGG 293
```

```
Qy     360 GGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGA 419
           |    |||||   |      ||| |  | || |  || ||  || |||  |||||
Db     294 TCTGTTGGCGGAGACATCGGCGAATGCCACGGTGCTAACGATTACGGCCTTCACGGTGGA 353

Qy     420 GCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGC 479
           ||| ||  | |||||    || ||||| |  | | |      ||| |  |   | || ||
Db     354 GCGGTATATTGCCATTTGCCATCCGTTTCTGGGCCAGGCCATGAGTAAACTCAGTCGCGC 413

Qy     480 CCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAG 539
           | |  | ||| ||| | | || |||    |  || |  |     || | ||  |  |
Db     414 CATTCGCATCATCGTCCTGGTTTGGATTATGGCCATAGTTACGGCCATTCCGCAGGCTGC 473

Qy     540 CATCCATGGCATCAAGTTCCACT 562
           |     ||| ||| ||  | | |
Db     474 CCAATTTGGAATCGAGCACTATT 496


RESULT 2
US-09-967-768A-296
; Sequence 296, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
;  APPLICANT: Augustus, Meena
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-72
;  CURRENT APPLICATION NUMBER: US/09/967,768A
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,109
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,034
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,111
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 296
;   LENGTH: 1427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-967-768A-296

  Query Match             12.8%;  Score 93;  DB 10;  Length 1427;
  Best Local Similarity   53.5%;  Pred. No. 1.4e-14;
  Matches  220;  Conservative    0;  Mismatches  185;  Indels    6;  Gaps    1;

Qy     143 TGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTC 202
           | || | ||   | |   |||||||| |  | ||||| | ||||||   | | ||||| |
Db     256 TCTACGCGCTGGTGTGCCTGGTGGGGCTGGTGGGCAACGCCCTGGTCATCTTCGTGATCC 315

Qy     203 TGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTG 262
           | |   ||      |||||||| | ||||||  ||||||  ||| ||||||  ||  | |
       316 TTCGCTACGCCAAGATGAAGACGGCTACCAACATCTACCTGCTCAACCTGGCCGTAGCCG 375

Qy     263 ACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACC 322
           ||   ||  || ||||    |       | |||     |  |       |||| |||  |
Db     376 ACGAGCTCTTCATGCTGAGCGTGCCCTTCGTGGCCTCGTCGGCCGCCCTGCGCCACTGGC 435

Qy     323 CTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCT 382
           |       |||||  |||||   |||  |     |     | ||  |    |       |
Db     436 C------CTTCGGCTCCGTGCTGTGCCGCGCGGTGCTCAGCGTCGACGGCCTCAACATGT 489

Qy     383 TCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACC 442
           || ||  | || || |  |||||    |||||||||| ||||||||||| |  | ||||
Db     490 TCACCAGCGTCTTCTGTCTCACCGTGCTCAGCGTGGACCGCTACGTGGCCGTGGTGCACC 549

Qy     443 CGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCT 502
           |  | |||||        |  |    |   ||  ||||      |||  |     ||| |
Db     550 CTCTGCGCGCGGCGACCTACCGGCGGCCCAGCGTGGCCAAGCTCATCAACCTGGGCGTGT 609

Qy     503 GGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCATCA 553
           ||      ||| || |  || | || |||| | ||| | ||   | || ||

Db     610 GGCTGGCATCCCTGTTGGTCACTCTCCCCATCGCCATCTTCGCAGACACCA 660


RESULT 3
US-09-970-966-208
; Sequence 208, Application US/09970966
; Patent No. US20020173638A1
; GENERAL INFORMATION:
;  APPLICANT: Stolk, John A.
;  APPLICANT:  Molesh, David Alan
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.484C6
;  CURRENT APPLICATION NUMBER: US/09/970,966
;  CURRENT FILING DATE:  2001-10-02
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 208
;   LENGTH: 1362
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-970-966-208

  Query Match             12.3%;  Score 89.4;  DB 9;  Length 1362;
  Best Local Similarity   54.4%;  Pred. No. 1.1e-13;
  Matches  202;  Conservative    0;  Mismatches  166;  Indels    3;  Gaps    1;

Qy     164 TGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGA 223
           |||||  ||  | ||   | | |||   || |||| | | | || |         || |||
Db     137 TGGGGAACAGCGCCACCATTCGGGTCACCCAGGTGCTGCAGAAGAAAGGATACTTGCAGA 196

Qy     224 CGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTG 283
           |       |   || || |  | ||  ||||   ||| ||| || |||| | ||| | |
Db     197 AGGAGGTGACAGACCACATGGTGAGTTTGGCTTGCTCGGACATCTTGGTGTTCCTCATCG 256

Qy     284 GAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGC---CCG 340
           | |||||| |||||  ||||    ||      | ||   ||     |  || | |     ||
Db     257 GCATGCCCATGGAGTTCTACAGCATCATCTGGAATCCCCTGACCACGTCCAGCTACACCC 316

Qy     341 TGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCA 400
           ||  |||| |  |  | ||   |||||| ||| ||     ||| |||  |  | ||   |
Db     317 TGTCCTGCAAGCTGCACACTTTCCTCTTCGAGGCCTGCAGCTACGCTACGCTGCTGCACG 376

Qy     401 TCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGC 460
           |    ||  ||||| | |||||||| | ||||||   ||||| ||| |   |||
Db     377 TGCTGACACTCAGCTTTGAGCGCTACATCGCCATCTGTCACCCCTTCAGGTACAAGGCTG 436

Qy     461 AGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCT 520
           |       |  ||| ||      | |  | ||| |||||||||  |  |||||  ||
Db     437 TGTCGGGACCTTGCCAGGTGAAGCTGCTGATTGGCTTCGTCTGGGTCACCTCCGCCCTGG 496

Qy     521 TCTCCCTGCCC 531
           |  | ||||||
Db     497 TGGCACTGCCC 507


RESULT 4
US-09-825-294-208
; Sequence 208, Application US/09825294
; Patent No. US20020004491A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.484C5
;  CURRENT APPLICATION NUMBER: US/09/825,294
;  CURRENT FILING DATE:  2001-04-03
;  NUMBER OF SEQ ID NOS: 215
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 208
;   LENGTH: 1362
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-825-294-208

  Query Match             12.3%;  Score 89.4;  DB 10;  Length 1362;
  Best Local Similarity   54.4%;  Pred. No. 1.1e-13;
  Matches  202;  Conservative    0;  Mismatches  166;  Indels    3;  Gaps    1;

Qy    164 TGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGA 223
          |||||  ||  | ||   | | |||   || |||| | | | || |        || |||
Db    137 TGGGGAACAGCGCCACCATTCGGGTCACCCAGGTGCTGCAGAAGAAAGGATACTTGCAGA 196

      224 CGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTG 283
           |       |   || || |  | ||  ||||   ||| ||| || ||||  | ||| | |
Db    197 AGGAGGTGACAGACCACATGGTGAGTTTGGCTTGCTCGGACATCTTGGTGTTCCTCATCG 256

Qy    284 GAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGC---CCG 340
          | |||||| ||||| ||||    ||     | ||   ||    |  || | |     ||
Db    257 GCATGCCCATGGAGTTCTACAGCATCATCTGGAATCCCCTGACCACGTCCAGCTACACCC 316

Qy    341 TGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCA 400
          ||  |||| |  |  | ||   |||||| ||| ||    ||| ||| |  | ||   |
Db    317 TGTCCTGCAAGCTGCACACTTTCCTCTTCGAGGCCTGCAGCTACGCTACGCTGCTGCACG 376

Qy    401 TCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGC 460
          |    ||  ||||| | ||||||||| | ||||||   ||||| ||| |   |||
Db    377 TGCTGACACTCAGCTTTGAGCGCTACATCGCCATCTGTCACCCCTTCAGGTACAAGGCTG 436

Qy    461 AGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCT 520
           |       |   ||| ||       | |  | ||| |||||||||  |  |||||  ||
Db    437 TGTCGGGACCTTGCCAGGTGAAGCTGCTGATTGGCTTCGTCTGGGTCACCTCCGCCCTGG 496

Qy    521 TCTCCCTGCCC 531
          |  | ||||||
Db    497 TGGCACTGCCC 507


RESULT 5
US-09-829-631A-12
; Sequence 12, Application US/09829631A
; Patent No. US20020091235A1
; GENERAL INFORMATION:
;  APPLICANT: Sibley, David R.
;  APPLICANT:  Monsma, Frederick J.
;  APPLICANT:  Hamblin, Mark
;  TITLE OF INVENTION: The ST-B17 Serotonin Receptor
;  FILE REFERENCE: NIH047.1CP1C1
;  CURRENT APPLICATION NUMBER: US/09/829,631A
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: US 08/428,242
;  PRIOR FILING DATE: 1995-09-18
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 1647
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (135)...(1454)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1647)
;   OTHER INFORMATION: n = A,T,C or G
US-09-829-631A-12

  Query Match             11.7%;  Score 85.6;  DB 10;  Length 1647;
  Best Local Similarity   53.8%;  Pred. No. 1.1e-12;
  Matches  200;  Conservative    0;  Mismatches  169;  Indels    3;  Gaps    1;

Qy    158 TTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTA 217
          | | | ||  ||     | |||    ||| ||  |  | | ||       || |  ||
Db    241 TCGCGCTGACGGCGGCGGCCAACTCGCTGCTGATCGCGCTCATCTGCACTCAGCCCGCGC 300

Qy    218 TGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGC 277
          ||   |   | ||||| || |||  |    ||       ||||||||  |||||  ||
Db    301 TGCGCAACACGTCCAACTTCTTCCTGGTGTCGCTCTTCACGTCTGACCTGATGGTCGGGC 360

Qy    278 TCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGC 337
          |  | | |||||   |  || |     ||    ||   ||   |  |   || | | ||
Db    361 TGGTGGTGATGCC---GCCGGCCATGCTGAACGCGCTGTACGGGCGCTGGGTGCTGGCGC 417

Qy    338 CCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCA 397
           ||    ||||    ||  ||| ||| ||   | ||     |||  ||||||||||||||
Db    418 GCGGCCTCTGCCTGCTCTGGACCGCCTTCGACGTGATGTGCTGCAGCGCCTCCATCCTCA 477

Qy    398 GCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAAC 457
          | ||  |  | ||||| |||| |||||| ||  ||||||    ||| | |||  ||| |
Db    478 ACCTCTGCCTCATCAGCCTGGACCGCTACCTGCTCATCCTCTCGCCGCTGCGCTACAAGC 537

Qy    458 TGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGC 517
          |||  |  || | |  || |||||     |  ||     || |||| || || ||| |
Db    538 TGCGCATGACGCCCCTGCGTGCCCTGGCCCTAGTCCTGGGCGGCTGGAGCCTCGCCGCTC 597

Qy    518 TCTTCTCCCTGC 529
          ||  |||| | |
Db    598 TCGCCTCCTTCC 609


RESULT 6
US-09-917-800A-1620
; Sequence 1620, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1620
;   LENGTH: 2108
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_016991
US-09-917-800A-1620

  Query Match             11.3%;  Score 82.6;  DB 10;  Length 2108;
  Best Local Similarity   47.5%;  Pred. No. 7e-12;
```

Matches 278; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

```
Qy    126 CCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCT 185
Db    149 CATCTCTGTGGGCCTGGTGCTGGGCGCCTTCATCCTCTTTGCCATCGTGGGCAACATCTT 208

Qy    186 GGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTT 245
Db    209 GGTCATCCTGTCGGTGGCCTGCAACCGGCACCTGCGGACGCCCACCAACTACTTTATCGT 268

Qy    246 CAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGA 305
Db    269 CAACCTGGCCATTGCTGACCTGCTGTTGAGTTTCACAGTACTGCCCTTCTCCGCTACCCT 328

Qy    306 GATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCT 365
Db    329 AGAAGTGCTTGGCTACTGGGTGCTGTTGAGTTTCTT---CTGTGACATCTGGGCAGCGGT 385

Qy    366 CTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTA 425
Db    386 AGATGTCCTGTGCTGTACGGCCTCCATCCTGAGCCTATGTGCCATCTCCATTGACCGCTA 445

Qy    426 CGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAG 485
Db    446 CATTGGGGTGCGATACTCTCTGCAGTACCCCACGCTGGTCACCCGCAGGAAGGCCATCTT 505

Qy    486 GATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCA 545
Db    506 GGCGCTCCTCAGTGTGTGGGTCTTGTCCACGGTCATCTCCATCGGGCCTCTCCTTGGATG 565

Qy    546 TGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTAC 605
Db    566 GAAAGAACCTGCGCCCAATGATGACAAAGAATGTGGGGTCACCGAAGAACCCTTCTACGC 625

Qy    606 GGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTA 665
Db    626 CCTCTTTTCCTCCCTGGGCTCCTTCTACATCCCGCTCGCGGTCATCCTGGTCATGTACTG 685

Qy    666 CCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGC 710
Db    686 CCGGGTCTACATCGTGGCCAAGAGGACCACCAAGAATCTGGAGGC 730
```

RESULT 7
US-10-077-870-1
; Sequence 1, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
; APPLICANT: Salonen, Jukka T
; TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
; FILE REFERENCE: 0933-0183P
; CURRENT APPLICATION NUMBER: US/10/077,870
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: FI 20010323
PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1341)
; OTHER INFORMATION: Coding sequence for variant human alpha-2B-adrenoceptor protein
US-10-077-870-1

Query Match 11.2%; Score 82; DB 9; Length 1344;
Best Local Similarity 52.3%; Pred. No. 8.5e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```
Qy    170 TCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCA 229
Db     80 TCTTCGGCAACGCTCTGGTCATCCTGGCTGTGTTGACCAGCCGCTCGCTGCGCGCCCCTC 139

Qy    230 CCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGC 289
Db    140 AGAACCTGTTCCTGGTGTCGCTGGCCGCCGCCGACATCCTGGTGGCCACGCTCATCATCC 199

Qy    290 CCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCT 349
Db    200 CTTTCTCGCTGGCCAACGAGCTGCTGGGCTAC---TGGTACTTCCGGCGCACGTGGTGCG 256

Qy    350 ACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCG 409
Db    257 AGGTGTACCTGGCGCTCGACGTGCTCTTCTGCACCTCGTCCATCGTGCACCTGTGCGCCA 316

Qy    410 TCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCC 469
Db    317 TCAGCCTGGACCGCTACTGGGCCGTGAGCCGCGCGCTGGAGTACAACTCCAAGCGCACCC 376

Qy    470 GGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGC 529
Db    377 CGCGCCGCATCAAGTGCATCATCCTCACTGTGTGGCTCATCGCCGCCGTCATCTCGCTGC 436

Qy    530 CCAACACCAGCATCCATGGCATCAAGTTCCACTA 563
Db    437 CGCCCCTCATCTACAAGGGCGACCAGGGCCCCCA 470
```

RESULT 8
US-09-825-923-1
; Sequence 1, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku
; APPLICANT: Pesonen, Ullamari
; APPLICANT: Scheinin, Mika
; APPLICANT: Salonen, Jukka T
; APPLICANT: Tuomainen, Tomi-Pekka
; APPLICANT: Lakka, Timo A
; APPLICANT: Nyyss"nen, Kristiina
; APPLICANT: Salonen, Riitta
; APPLICANT: Kauhanen, Jussi
; APPLICANT: Valkonen, Veli-Pekka
; TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
; TITLE OF INVENTION: protein, and uses thereof
; FILE REFERENCE: Alpha-2B-AR variant
; CURRENT APPLICATION NUMBER: US/09/825,923
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/422,985
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1341)
; OTHER INFORMATION: Coding sequence for variant human
; OTHER INFORMATION: alpha-2B-adrenoceptor protein
US-09-825-923-1

Query Match 11.2%; Score 82; DB 10; Length 1344;
Best Local Similarity 52.3%; Pred. No. 8.5e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```
Qy    170 TCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCA 229
Db     80 TCTTCGGCAACGCTCTGGTCATCCTGGCTGTGTTGACCAGCCGCTCGCTGCGCGCCCCTC 139
```

```
Qy    230 CCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGC 289
Db    140 AGAACCTGTTCCTGGTGTCGCTGGCCGCCGCCGACATCCTGGTGGCCACGCTCATCATCC 199

Qy    290 CCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCT 349
Db    200 CTTTCTCGCTGGCCAACGAGCTGCTGGGCTAC---TGGTACTTCCGGCGCACGTGGTGCG 256

Qy    350 ACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCG 409
Db    257 AGGTGTACCTGGCGCTCGACGTGCTCTTCTGCACCTCGTCCATCGTGCACCTGTGCGCCA 316

Qy    410 TCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCC 469
Db    317 TCAGCCTGGACCGCTACTGGGCCGTGAGCCGCGCGCTGGAGTACAACTCCAAGCGCACCC 376

Qy    470 GGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGC 529
Db    377 CGCGCCGCATCAAGTGCATCATCCTCACTGTGTGGCTCATCGCCGCCGTCATCTCGCTGC 436

Qy    530 CCAACACCAGCATCCATGGCATCAAGTTCCACTA 563
Db    437 CGCCCCTCATCTACAAGGGCGACCAGGGCCCCCA 470
```

RESULT 9
US-10-077-870-3
; Sequence 3, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
; APPLICANT: Salonen, Jukka T
; TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
; FILE REFERENCE: 0933-0183P
; CURRENT APPLICATION NUMBER: US/10/077,870
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: FI 20010323
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
, NAME/KEY: CDS
; LOCATION: (1)..(1350)
; OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor protein
US-10-077-870-3

Query Match 11.2%; Score 82; DB 9; Length 1353;
Best Local Similarity 52.3%; Pred. No. 8.6e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```
Qy    170 TCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCA 229
Db     80 TCTTCGGCAACGCTCTGGTCATCCTGGCTGTGTTGACCAGCCGCTCGCTGCGCGCCCCTC 139

Qy    230 CCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGC 289
Db    140 AGAACCTGTTCCTGGTGTCGCTGGCCGCCGCCGACATCCTGGTGGCCACGCTCATCATCC 199

Qy    290 CCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCT 349
Db    200 CTTTCTCGCTGGCCAACGAGCTGCTGGGCTAC---TGGTACTTCCGGCGCACGTGGTGCG 256

Qy    350 ACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCG 409
Db    257 AGGTGTACCTGGCGCTCGACGTGCTCTTCTGCACCTCGTCCATCGTGCACCTGTGCGCCA 316

Qy    410 TCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCC 469
Db    317 TCAGCCTGGACCGCTACTGGGCCGTGAGCCGCGCGCTGGAGTACAACTCCAAGCGCACCC 376

Qy    470 GGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGC 529
Db    377 CGCGCCGCATCAAGTGCATCATCCTCACTGTGTGGCTCATCGCCGCCGTCATCTCGCTGC 436

Qy    530 CCAACACCAGCATCCATGGCATCAAGTTCCACTA 563
Db    437 CGCCCCTCATCTACAAGGGCGACCAGGGCCCCCA 470
```

RESULT 10
US-09-825-923-3
; Sequence 3, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku
; APPLICANT: Pesonen, Ullamari
; APPLICANT: Scheinin, Mika
; APPLICANT: Salonen, Jukka T
; APPLICANT: Tuomainen, Tomi-Pekka
; APPLICANT: Lakka, Timo A
; APPLICANT: Nyyss"nen, Kristiina
; APPLICANT: Salonen, Riitta
; APPLICANT: Kauhanen, Jussi
; APPLICANT: Valkonen, Veli-Pekka
; TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
; TITLE OF INVENTION: protein, and uses thereof
; FILE REFERENCE: Alpha-2B-AR variant
; CURRENT APPLICATION NUMBER: US/09/825,923
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/422,985
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1350)
; OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor
; OTHER INFORMATION: protein
US-09-825-923-3

Query Match 11.2%; Score 82; DB 10; Length 1353;
Best Local Similarity 52.3%; Pred. No. 8.6e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```
Qy    170 TCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCA 229
Db     80 TCTTCGGCAACGCTCTGGTCATCCTGGCTGTGTTGACCAGCCGCTCGCTGCGCGCCCCTC 139

Qy    230 CCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGC 289
Db    140 AGAACCTGTTCCTGGTGTCGCTGGCCGCCGCCGACATCCTGGTGGCCACGCTCATCATCC 199

Qy    290 CCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCT 349
Db    200 CTTTCTCGCTGGCCAACGAGCTGCTGGGCTAC---TGGTACTTCCGGCGCACGTGGTGCG 256

Qy    350 ACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCG 409
Db    257 AGGTGTACCTGGCGCTCGACGTGCTCTTCTGCACCTCGTCCATCGTGCACCTGTGCGCCA 316

Qy    410 TCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCC 469
Db    317 TCAGCCTGGACCGCTACTGGGCCGTGAGCCGCGCGCTGGAGTACAACTCCAAGCGCACCC 376
```

```
Qy    470 GGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGC 529
          ||||||   |    | ||| ||  ||  || ||| | || |||   || |||| ||||
Db    377 CGCGCCGCATCAAGTGCATCATCCTCACTGTGTGGCTCATCGCCGCCGTCATCTCGCTGC 436

Qy    530 CCAACACCAGCATCCATGGCATCAAGTTCCACTA 563
          |   |  || |  | | |||  | ||  || | |
Db    437 CGCCCCTCATCTACAAGGGCGACCAGGGCCCCCA 470


RESULT 11
US-10-044-090-444
; Sequence 444, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 444
;   LENGTH: 2487
;   TYPE: DNA
    ORGANISM: Homo sapiens
    FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1704944CB1
US-10-044-090-444

  Query Match             10.8%;  Score 79;  DB 12;  Length 2487;
  Best Local Similarity   52.2%;  Pred. No. 6.2e-11;
  Matches  206;  Conservative    0;  Mismatches  180;  Indels    9;  Gaps    1;

Qy    141 GGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGAT 200
          |||| |  || | |||  |||| |||   || | |||||||| |||||||||  ||   |
Db    415 GGTGGAGCTGGCCATTGCTGTGCTGGCCATCCTGGGCAATGTGCTGGTGTGCTGGGCCGT 474

Qy    201 TCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTC 260
            ||  ||       || |||    ||||||||||   |  |    |||||||   |
Db    475 GTGGCTCAACAGCAACCTGCAGAACGTCACCAACTACTTTGTGGTGTCACTGGCGGCGGC 534

Qy    261 TGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTA 320
          ||| ||   ||     | || |  || ||| |          ||  ||      || | |
Db    535 CGACATCGCAGTGGGTGTGCTCGCCATCCCCTT---------TGCCATCACCATCAGCAA 585

Qy    321 CCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTG 380
          |   ||||     |    |   ||||||  |||||   |   | || |   |  | |
Db    586 CGGGTTCTGCGCTGCCTGCCACGGCTGCCTCTTCATTGCCTGCTTCGTCCTGGTCCTCAC 645

Qy    381 CTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACA 440
          |||||| ||||  ||    || ||  | | || |||||| | |||||||
      646 GCAGAGCTCCATCTTCAGTCTCCTGGCCATCGCCATTGACCGCTACATTGCCATCCGCAT 705

      441 CCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGT 500
          |||| ||||   |||      |   ||| |     ||||     | ||| | | ||||
Db    706 CCCGCTCCGGTACAATGGCTTGGTGACCGGCACGAGGGCTAAGGGCATCATTGCCATCTG 765

Qy    501 CTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACA 535
          |||||   | ||  |   | ||  |||| |   ||
Db    766 CTGGGTGCTGTCGTTTGCCATCGGCCTGACTCCCA 800


RESULT 12
US-10-052-589-1
; Sequence 1, Application US/10052589
; Patent No. US20020133832A1
; GENERAL INFORMATION:
;  APPLICANT: Perez, Dianne
;  APPLICANT:  Zuscik, Michael
;  TITLE OF INVENTION: Model systems for neurodegenerative and cardiovascular disorders
;  FILE REFERENCE: 26473/04200
;  CURRENT APPLICATION NUMBER: US/10/052,589
;  CURRENT FILING DATE:  2002-01-18
;  PRIOR APPLICATION NUMBER: US 09/568,255
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 2048
;   TYPE: DNA
;   ORGANISM: Mesocricetus sp.
US-10-052-589-1

  Query Match             10.7%;  Score 77.8;  DB 12;  Length 2048;
  Best Local Similarity   47.0%;  Pred. No. 1.2e-10;
  Matches  275;  Conservative    0;  Mismatches  307;  Indels    3;  Gaps    1;

Qy    126 CCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCT 185
          | || | |||   |||||   |   |  |  |  |  | |   |  | |||||  ||||
Db    149 CATCTCTGTGGGCCTGGTGCTGGGCGCCTTCATCCTCTTTGCCATTGTGGGCAACATCCT 208

Qy    186 GGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTT 245
          |||   ||||    |         | | |    ||  ||||||||||||||||| | |  |
Db    209 GGTCATCCTGTCAGTGGCCTGCAATCGGCACCTGCGGACGCCCACCAACTACTTCATTGT 268

Qy    246 CAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGA 305
          || ||||||  |  |||||||| ||| |     ||   |   |||||  |    |
Db    269 CAACCTGGCCATTGCTGACCTGCTGTTGAGTTTCACAGTCCTGCCCTTCTCCGCTACCCT 328

Qy    306 GATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCT 365
                ||    ||||    |  ||   |||| | |   |||  || ||  | | ||  |
Db    329 AGAAGTGCTTGGCTACTGGGTTCTG---GGGCGCATCTTCTGTGACATCTGGGCAGCGGT 385

Qy    366 CTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTA 425
              |         ||    |||||||||||| ||| |      || ||  | | || |||||
Db    386 GGACGTCCTGTGCTGTACGGCCTCCATCCTGAGCCTATGTGCCATCTCCATTGATCGCTA 445

Qy    426 CGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAG 485
          | | |   | |   || | | |     |       | |  ||||||  |   |||| ||
Db    446 CATTGGGGTGCGCTACTCTCTGCAGTACCCCACTCTGGTCACCCGCAGGAAGGCCATCTT 505

Qy    486 GATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCA 545
          |   |||  ||  || |||||  || |||  | || ||||| |          |
Db    506 GGCACTCCTCAGTGTGTGGGTTTTGTCCACGGTCATCTCCATCGGGCCTCTCCTTGGATG 565

Qy    546 TGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTAC 605
                    | | ||  |  ||| |  | |   ||| | |       ||  || | |
Db    566 GAAAGAACCAGCGCCCAACGACGACAAGGAATGCGGAGTCACCGAAGAACCCTTCTATGC 625

Qy    606 GGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTA 665
            || |     || || |   || | |  ||     ||  |||||       || | | ||
Db    626 CCTCTTTTCCTCCCTGGGCTCCTTCTACATCCCACTCGCGGTCATTCTGGTCATGTACTG 685

Qy    666 CCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGC 710
          ||   ||  |||    | ||  ||   |  | |  ||||   |||
Db    686 CCGGGTCTACATCGTGGCCAAGAGGACCACCAAGAACCTGGAGGC 730


RESULT 13
US-09-964-824A-249
; Sequence 249, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
;  APPLICANT: Horrigan, Stephen
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-73
;  CURRENT APPLICATION NUMBER: US/09/964,824A
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: US/60/236,033
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,032
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,028
;  PRIOR FILING DATE: 2000-09-28
```

```
;  NUMBER OF SEQ ID NOS: 583
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 249
;   LENGTH: 2595
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2595)
;   OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-249

  Query Match            10.7%;  Score 77.8;  DB 10;  Length 2595;
  Best Local Similarity  52.3%;  Pred. No. 1.3e-10;
  Matches  197;  Conservative    0;  Mismatches  177;  Indels    3;  Gaps    1;

Qy     161 TGGTGGGGGTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGA 220
Db     931 TGGTGACTGTCGTGGGCAACATCCTGGTGATGCTGTCCATCAAGGTCAACAGGCAGCTGC 990

Qy     221 AGACGCCCACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCC 280
Db     991 AGACAGTCAACAACTACTTCCTCTTCAGCCTGGCGTGTGCTGATCTCATCATAGGCGCCT 1050

Qy     281 TTGGAATGCCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCG 340
Db    1051 TCTCCATGAACCTCTACACCGTGTACATCATCAAGGGCTACTGGCCCCTG---GGCGCCG 1107

Qy     341 TGGGCTGCTACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCA 400
Db    1108 TGGTCTGCGACCTGTGGCTGGCCCTGGACTACGTGGTGAGCAACGCCTCCGTCATGAACC 1167

Qy     401 TCACCACCGTCAGCGTGGAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGC 460
Db    1168 TTCTCATCATCAGCTTTGACCGCTACTTCTGCGTCACCAAGCCTCTCACCTACCCTGCCC 1227

Qy     461 AGAGCACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCT 520
Db    1228 GGCGCACCACCAAGATGGCAGGCCTCATGATTGCTGCTGCCTGGGTACTGTCCTTCGTGC 1287

Qy     521 TCTCCCTGCCCAACACC 537
Db    1288 TCTGGGCGCCTGCCATC 1304


RESULT 14
US-09-940-240-15
; Sequence 15, Application US/09940240
; Patent No. US20020166133A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  TITLE OF INVENTION: NEOKINE PROTEIN AND NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: MBIO98-008CP1CN1M
;  CURRENT APPLICATION NUMBER: US/09/940,240
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/248,239
;  PRIOR FILING DATE: 1999-02-10
;  PRIOR APPLICATION NUMBER: 09/023,664
;  PRIOR FILING DATE: 1998-02-08
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 2050
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)..(1176)
US-09-940-240-15

  Query Match            10.6%;  Score 77;  DB 9;  Length 2050;
  Best Local Similarity  47.3%;  Pred. No. 1.9e-10;
  Matches  304;  Conservative    0;  Mismatches  330;  Indels    9;  Gaps    2;

Qy      72 CAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCACTTCTTCCTCCC 131
Db     168 CATCGTCGTGGACACCGTGCTGTGCCCCAACATGCCCAACAAAAGCGTGCTGCTGTACAC 227

Qy     132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
Db     228 GCTGTCCTTCATTTACATCTTCATCTTCGTGATCGGCATGATCGCCAACTCCGTGGTGGT 287

Qy     192 CCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTACCTCTTCAGCCT 251
Db     288 CTGGGTGAACATCCAGGCCAAGACCACCGGCTACGACACTCACTGCTACATCCTCAACCT 347

Qy     252 GGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTCTATGAGATGTG 311
Db     348 GGCCATCGCCGACCTGTGGGTGGTCGTCACCATCCCCGTCTGGGTGGTCAGCCTCGTGCA 407

Qy     312 GCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACGGCCCTCTTTGA 371
Db     408 GCATAACCAGTGGCCCATGGGGGAGCTCACG---TGCAAGATCACGCACCTCATCTTCTC 464

Qy     372 GACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAGCGCTACGTGGC 431
Db     465 CATCAACCTGTTCGGCAGCATCTTCTTCCTCACGTGCATGAGCGTGGACCGCTACCTCTC 524

Qy     432 CATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCCCTCAGGATCCT 491
Db     525 CAT------CACCTACTTCGCCAGCACGTCGAGCCGCAGGAAGAAGGTGGTTCGCCGCGC 578

Qy     492 CGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGCATCCATGGCAT 551
Db     579 CGTCTGTGTCCTGGTGTGGCTGCTGGCCTTCTGCGTGTCCCTGCCCGACACCTACTACCT 638

Qy     552 CAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACCTGTACGGTCAT 611
Db     639 GAAGACCGTCACGTCGGCGTCCAACAACGAGACCTACTGCCGCTCCTTCTACCCCGAGCA 698

Qy     612 CAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTATTCTACCTCCT 671
Db     699 CAGCGTCAAGGAGTGGCTCATCAGCATGGAGCTGGTCTCGGTGGTCCTGGGCTTCGCCAT 758

Qy     672 CCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTC 714
Db     759 CCCCTTCTGCGTCATCGCCGTCTTCTACTGCCTGCTGGCCCGC 801


RESULT 15
US-09-954-531-995
; Sequence 995, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
;  APPLICANT: Weaver, Zoe
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
;  TITLE OF INVENTION:  Gene Sets
;  FILE REFERENCE: 689290-77
;  CURRENT APPLICATION NUMBER: US/09/954,531
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: US/60/233,133
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,009
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,034
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,509
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: US/60/234,567
;  PRIOR FILING DATE: 2000-09-22
;  NUMBER OF SEQ ID NOS: 1392
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 995
;   LENGTH: 2625
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-09-954-531-995

  Query Match            10.6%;  Score 77;  DB 9;  Length 2625;
  Best Local Similarity  49.0%;  Pred. No. 2e-10;
  Matches  265;  Conservative    0;  Mismatches  270;  Indels    6;  Gaps    2;

Qy     169 GTCATTGGCAATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCC 228
           ||| | ||||| || |||||||||| |||    |    |   |  ||||  || ||||  ||
Db     310 GTCTTCGGCAACGTGCTGGTGTGCATGGCTGTGTCCCGCGAGAAGGCGCTGCAGACCACC 369

Qy     229 ACCAACTACTACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATG 288
           |||||||||    || |||||||  || ||  | ||||||||  |||       || |  |||
Db     370 ACCAACTACCTGATCGTCAGCCTCGCAGTGGCCGACCTCCTCGTCGCCACACTGGTCATG 429

Qy     289 CCCCTGGAGGTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGC 348
           |||  ||  |||||   |  |  |        |    |     ||| |     |   |||
Db     430 CCCTGGGTTGTCTACCTGGAGGTGGTAGGTGA---GTGGAAATTCAGCAGGATTCACTGT 486

Qy     349 TACTTCAAGACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACC 408
           || ||       | ||    |  |   |||||   ||   |||||| | | |     ||
Db     487 GACATCTTCGTCACTCTGGACGTCATGATGTGCACGGCGAGCATCCTGAACTTGTGTGCC 546

Qy     409 GTCAGCGTGGAGCGCTACGTGGCCAT---CCTACACCCGTTCCGCGCCAAACTGCAGAGC 465
           |||||  | ||  | |||   ||  |   | | | |  | |   |   |  |   | |||
Db     547 ATCAGCATCGACAGGTACACAGCTGTGGCCATGCCCATGCTGTACAATACGCGCTACAGC 606

Qy     466 ACCCGGCGCCGGGCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCC 525
           ||  |||||||| |  |   ||  ||  |||||||||| | | ||| |   | |||||
Db     607 TCCAAGCGCCGGGTCACCGTCATGATCTCCATCGTCTGGGTCCTGTCCTTCACCATCTCC 666

Qy     526 CTGCCCAACACCAGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTC 585
             ||   |  |  |   |     | |     |||  ||     ||      ||
Db     667 TGCCCACTCCTCTTCGGACTCAATAACGCAGACCAGAACGAGTGCATCATTGCCAACCCG 726

Qy     586 CCAGGTTCGGCCACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAG 645
           |      || |  ||   |  || ||     |     |  || || | ||| | | | |
Db     727 GCCTTCGTGGTCTACTCCTCCATCGTCTCCTTCTACGTGCCCTTCATTGTCACCCTGCTG 786

Qy     646 GTCACCTCCTTCCTATTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTC 705
           |||  |  |    | | |    ||||||  ||       ||   | |||  |  | | | |
Db     787 GTCTACATCAAGATCTACATTGTCCTCCGCAGACGCCGCAAGCGAGTCAACACCAAACGC 846

Qy     706 A 706
           |
Db     847 A 847


Search completed: January 17, 2003, 02:36:32
Job time : 70 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 03:44:22 ; Search time 82 Seconds
(without alignments)
608.090 Million cell updates/sec

Title: US-09-684-725-2
Perfect score: 1263
Sequence: 1 MEKLQNASWIYQQKLEDPFQ.....LLPMTVISVLYYLMALRVSI 242

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1252	99.1	415	4	Q9GZQ4	Q9gzq4 homo sapien
2	1241	98.3	415	4	Q96AM5	Q96am5 homo sapien
3	1039.5	82.3	395	11	Q91Z76	Q91z76 mus musculu
4	1033.5	81.8	395	11	Q9ESQ4	Q9esq4 rattus norv
5	1033.5	81.8	395	11	Q9JIB1	Q9jib1 rattus norv
6	680.5	53.9	426	4	Q9HB89	Q9hb89 homo sapien
7	642	50.8	402	11	Q9JIB2	Q9jib2 rattus norv
8	642	50.8	412	11	Q9JJI5	Q9jji5 rattus norv
9	596	47.2	405	11	O55040	O55040 mus musculu
10	418	33.1	428	5	Q9VFN4	Q9vfn4 drosophila
11	352	27.9	595	5	Q9VFW6	Q9vfw6 drosophila
12	337	26.7	660	5	Q9VFW5	Q9vfw5 drosophila
13	332.5	26.3	378	5	Q18701	Q18701 caenorhabdi
14	303.5	24.0	418	5	O17239	O17239 caenorhabdi
15	303	24.0	374	13	O93412	O93412 sphoeroides
16	301.5	23.9	363	13	O93413	O93413 sphoeroides
17	295	23.4	397	13	Q9DDR1	Q9ddr1 xenopus lae
18	290	23.0	559	13	O93414	O93414 sphoeroides
19	281	22.2	416	11	Q920Q5	Q920q5 mus musculu
20	277	21.9	416	11	Q8VIF5	Q8vif5 mus musculu
21	259	20.5	401	13	Q9DDR0	Q9ddr0 xenopus lae
22	257.5	20.4	400	6	Q95M54	Q95m54 macaca fasc
23	257	20.3	464	5	Q9GQ54	Q9gq54 aedes aegyp
24	256.5	20.3	377	13	Q98U14	Q98u14 brachydanio
25	256	20.3	410	4	Q8TBH6	Q8tbh6 homo sapien
26	255	20.2	390	11	Q8VI71	Q8vi71 mus musculu
27	255	20.2	391	11	Q8VI70	Q8vi70 mus musculu
28	255	20.2	393	11	Q9R1M0	Q9r1m0 mus musculu
29	255	20.2	401	11	Q9R1L9	Q9r1l9 mus musculu
30	255	20.2	409	11	Q8VI69	Q8vi69 mus musculu
31	255	20.2	438	11	Q9R0D1	Q9r0d1 mus musculu
32	255	20.2	444	11	Q9JIY1	Q9jiy1 mus musculu
33	253.5	20.1	454	4	Q9H573	Q9h573 homo sapien
34	253	20.0	383	13	O42324	O42324 catostomus
35	252.5	20.0	346	4	Q96GE0	Q96ge0 homo sapien
36	252.5	20.0	356	4	Q96TF2	Q96tf2 homo sapien
37	250.5	19.8	395	5	Q95YD7	Q95yd7 caenorhabdi
38	248	19.6	404	13	Q9DFB0	Q9dfb0 catostomus
39	248	19.6	431	5	Q8T8D1	Q8t8d1 urechis uni
40	246	19.5	513	13	Q9DFA9	Q9dfa9 catostomus
41	244	19.3	370	13	Q8UWL5	Q8uwl5 fugu rubrip
42	243.5	19.3	390	13	Q8QGQ4	Q8qgq4 carassius a
43	243	19.2	414	6	Q9GKA0	Q9gka0 canis famil
44	243	19.2	414	6	Q9GK99	Q9gk99 canis famil
45	243	19.2	443	6	Q9GJU1	Q9gju1 canis famil

ALIGNMENTS

RESULT 1
Q9GZQ4
ID Q9GZQ4 PRELIMINARY; PRT; 415 AA.
AC Q9GZQ4; Q9NRA6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Neuromedin U receptor 2 (Neuromedin U receptor-type 2) (G
DE protein-coupled receptor TGR-1).
GN NMUR2 OR NMU2R OR TGR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vaysse P.J., Branchek T.A., Gerald C., Forray C.,
RA Adham N.;
RT "Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System.";
RL J. Biol. Chem. 275:32452-32459(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11010960;
RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
RA Gustafson E.L., Monsma F.J. Jr., Hedrick J.A.;
RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in
RT the Central Nervous System.";
RL J. Biol. Chem. 275:39482-39486(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Pang L., Wang S., Laz T., Hedrick J.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE OF 4-415 FROM N.A.

```
RX   MEDLINE=20351041; PubMed=10894543;
RA   Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA   Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA   Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., Mckee K.K.,
RA   Hreniuk D.L., Mcdonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA   Caskey T., van der Ploeg L.H.T., Liu Q.;
RT   "Identification of receptors for neuromedin U and its role in
RT   feeding.";
RL   Nature 406:70-74(2000).
RN   [5]
RP   SEQUENCE OF 4-415 FROM N.A.
RX   PubMed=10887190;
RA   Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA   Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA   Fujino M.;
RT   "Identification and Functional Characterization of a Novel Subtype of
RT   Neuromedin U Receptor.";
RL   J. Biol. Chem. 275:29528-29532(2000).
DR   EMBL; AF272363; AAG24794.1; -.
DR   EMBL; AF292402; AAG03064.1; -.
DR   EMBL; AF242874; AAF82755.1; -.
DR   EMBL; AB041228; BAB13721.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01567; NEUROMEDNU2R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   415 AA;  47725 MW;  C2BACD84B313390F CRC64;

  Query Match             99.1%;  Score 1252;  DB 4;  Length 415;
  Best Local Similarity   99.6%;  Pred. No. 7e-109;
  Matches  239;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 2
Q96AM5
ID   Q96AM5      PRELIMINARY;      PRT;   415 AA.
AC   Q96AM5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neuromedin U receptor 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC016938; AAH16938.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01567; NEUROMEDNU2R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   415 AA;  47770 MW;  30BFEDD706436AB9 CRC64;

  Query Match             98.3%;  Score 1241;  DB 4;  Length 415;
  Best Local Similarity   99.2%;  Pred. No. 7.5e-108;
  Matches  238;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 63

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 123

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 183

Qy     181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||:
Db     184 IHGIKFHYFPNGSLVPGSATFTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRL 243


RESULT 3
Q91Z76
ID   Q91Z76      PRELIMINARY;      PRT;   395 AA.
AC   Q91Z76;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neuromedin U receptor type 2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,
RA   Gustafson E.L.;
RT   "Characterization of murine neuromedin U R2 receptor.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY057384; AAL26695.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01567; NEUROMEDNU2R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   395 AA;  44844 MW;  0BB540024566903D CRC64;

  Query Match             82.3%;  Score 1039.5;  DB 11;  Length 395;
  Best Local Similarity   81.7%;  Pred. No. 4.5e-89;
  Matches  196;  Conservative   19;  Mismatches   20;  Indels    5;  Gaps    1;

Qy       1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
           | ||:|||||:     |   |:||||||||:||||:||   |||||||  |||||||||
Db       1 MGKLENASWIH-----DSLMKYLNSTEEYLAYLCGPKRSDLSLPVSVVYALIFVVGVIGN 55

Qy      61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
           :|||||| :|| :|||||||||||||||||||||||||||||:| ||||||||||||||||
Db      56 LLVCLVIARHQTLKTPTNYYLFSLAVSDLLVLLLGMPLEVYELWHNYPFLFGPVGCYFKT 115

Qy     121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
```

```
          ||||||||||||||:||||:||||||:|||||||:||||||||| :|| |||:|||||||
Db    116 ALFETVCFASILSVTTVSIERYVAIVHPFRAKLESTRRRALRILSLVWSFSVVFSLPNTS 175

Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          ||||||  ||||| |||||||| ||||:|||||| ||||||:||||:|||||||| ||:
Db    176 IHGIKFQQFPNGSSVPGSATCTVTKPMWVYNFIIQATSFLFYILPMTLISVLYYLMGLRL 235


RESULT 4
Q9ESQ4
ID   Q9ESQ4      PRELIMINARY;      PRT;   395 AA.
AC   Q9ESQ4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor TGR-1.
GN   TGR-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20449029; PubMed=10887190;
RA   Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA   Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA   Fujino M.;
RT   "Identification and Functional Characterization of a Novel Subtype of
RT   Neuromedin U Receptor.";
RL   J. Biol. Chem. 275:29528-29532(2000).
DR   EMBL; AB041229; BAB13722.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01567; NEUROMEDNU2R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   395 AA;  44756 MW;  EABD6D36E9C355DA CRC64;

  Query Match           81.8%;  Score 1033.5;  DB 11;  Length 395;
  Best Local Similarity  80.4%;  Pred. No. 1.6e-88;
  Matches  193;  Conservative   22;  Mismatches   20;  Indels    5;  Gaps    1;

Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          | ||:|||||:     ||  |:||||||||| ||||:||   ||||| |  ||:|||:||
Db      1 MGKLENASWIH-----DPLMKYLNSTEEYLAHLCGPKRSDLSLPVSVAYALIFLVGVMGN 55

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          :|||:||::|| :|||||||||||||||||||||||:|||| ||||||||||||||
Db     56 LLVCMVIVRHQTLKTPTNYYLFSLAVSDLLVLLLGMPLEIYEMWHNYPFLFGPVGCYFKT 115

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          |||||||||||||:|||||||||||:||||||||:|||||||| :|| |||:|||||||
Db    116 ALFETVCFASILSVTTVSVERYVAIVHPFRAKLESTRRRALRILSLVWSFSVVFSLPNTS 175

Qy    181 IHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          |||||| :||||| |||||||| ||||:|| ||| ||||||:||||:|||||||| ||:
Db    176 IHGIKFQHFPNGSSVPGSATCTVTKPMWVYNLIIQATSFLFYILPMTLISVLYYLMGLRL 235


RESULT 5
Q9JIB1
ID   Q9JIB1      PRELIMINARY;      PRT;   395 AA.
AC   Q9JIB1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neuromedin U receptor 2.
GN   NMU2R.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=20351041; PubMed=10894543;
RA   Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA   Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA   Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., Mckee K.K.,
RA   Hreniuk D.L., Mcdonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA   Caskey T., van der Ploeg L.H.T., Liu Q.;
RT   "Identification of receptors for neuromedin U and its role in
RT   feeding.";
RL   Nature 406:70-74(2000).
DR   EMBL; AF242875; AAF82756.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01567; NEUROMEDNU2R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   395 AA;  44722 MW;  01D3765B5D5355C0 CRC64;

  Query Match           81.8%;  Score 1033.5;  DB 11;  Length 395;
  Best Local Similarity  80.4%;  Pred. No. 1.6e-88;
  Matches  193;  Conservative   22;  Mismatches   20;  Indels    5;  Gaps    1;

Qy      1 MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGN 60
          | ||:|||||:     ||  |:||||||||| ||||:||   ||||| |  ||:|||:||
Db      1 MGKLENASWIH-----DPLMKYLNSTEEYLAHLCGPKRSDLSLPVSVAYALIFLVGVMGN 55

Qy     61 VLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKT 120
          :|||:||::|| :|||||||||||||||||||||||:|||| ||||||||||||||
Db     56 LLVCMVIVRHQTLKTPTNYYLFSLAVSDLLVLLLGMPLEIYEMWHNYPFLFGPVGCYFKT 115

Qy    121 ALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTS 180
          |||||||||||||:|||||||||||:||||||||:|||||||| :|| |||:|||||||
Db    116 ALFETVCFASILSVTTVSVERYVAIVHPFRAKLESTRRRALRILSLVWSFSVVFSLPNTS 175

Qy    181 IHGIKFHYFPNGSSVPGSATCTVTKPMWVYNLIIQATSFLFYILPMTLISVLYYLMGLRL 240
          |||||| :||||| |||||||| ||||:|| ||| ||||||:||||:|||||||| ||:
Db    176 IHGIKFQHFPNGSSVPGSATCTVTKPMWVYNLIIQATSFLFYILPMTLISVLYYLMGLRL 235


RESULT 6
Q9HB89
ID   Q9HB89      PRELIMINARY;      PRT;   426 AA.
AC   Q9HB89; O43664;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neuromedin U receptor 1 (Orphan G protein-coupled receptor).
GN   NMUR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20490668; PubMed=10899166;
RA   Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA   Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA   Dai M., Lerman G.S., Vaysse P.J., Branchek T.A., Gerald C., Forray C.,
RA   Adham N.;
RT   "Identification and Characterization of Two Neuromedin U Receptors
RT   Differentially Expressed in Peripheral Tissues and the Central Nervous
RT   System.";
RL   J. Biol. Chem. 275:32452-32459(2000).
RN   [2]
```

```
RP   SEQUENCE OF 24-426 FROM N.A.
RX   MEDLINE=99000845; PubMed=9782091;
RA   Tan C.P., McKee K.K., Liu Q., Palyha O.C., Feighner S.D.,
RA   Hreniuk D.L., Smith R.G., Howard A.D.;
RT   "Cloning and characterization of a human and murine T-cell orphan G-
RT   protein-coupled receptor similar to the growth hormone secretagogue
RT   and neurotensin receptors.";
RL   Genomics 52:223-229(1998).
DR   EMBL; AF272362; AAG24793.1; -.
DR   EMBL; AF044601; AAC02680.1; -.
DR   EMBL; AF044600; AAC02680.1; JOINED.
DR   HSSP; P02699; 1F88.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01566; NEUROMEDNU1R.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   426 AA;  47350 MW;  F8DD781C020F04AA CRC64;

  Query Match             53.9%;  Score 680.5;  DB 4;  Length 426;
  Best Local Similarity   56.0%;  Pred. No. 1.5e-55;
  Matches  126;  Conservative   42;  Mismatches   54;  Indels    3;  Gaps    2;

Qy      17 DPFQKHLNSTEEYLAF-LCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKT 75
           ||  : || |:| |     ||:::  |:|:   |: ||||| :|| | ||||||:|:||:|
Db      36 DP--EDLNLTDEALRLKYLGPQQTELFMPICATYLLIFVVGAVGNGLTCLVILRHKAMRT 93

Qy      76 PTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSIT 135
           |||||||||||||||||:|:|||:|||| ||||| |  ||||:| ||| || ||:|:::|
Db      94 PTNYYLFSLAVSDLLVLLVGLPLELYEMWHNYPFLLGVGGCYFRTLLFEMVCLASVLNVT 153

Qy     136 TVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLV 195
           :|||||||::|| :|:   ||    |:|| ||| ::| ||||||:|||:  : |    |
Db     154 ALSVERYVAVVHPLQARSMVTRAHVRRVLGAVWGLAMLCSLPNTSLHGIRQLHVPCRGPV 213

Qy     196 PGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           | || | :::|  :|| ::| |: ||: ||| ::|||| |: ||:
Db     214 PDSAVCMLVRPRALYNMVVQTTALLFFCLPMAIMSVLYLLIGLRL 258


RESULT 7
Q9JIB2
ID   Q9JIB2      PRELIMINARY;      PRT;   402 AA.
AC   Q9JIB2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Neuromedin U receptor 1.
GN   NMU1R.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=20351041; PubMed=10894543;
RA   Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA   Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA   Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., Mckee K.K.,
RA   Hreniuk D.L., Mcdonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA   Caskey T., van der Ploeg L.H.T., Liu Q.;
RT   "Identification of receptors for neuromedin U and its role in
RT   feeding.";
RL   Nature 406:70-74(2000).
DR   EMBL; AF242873; AAF82754.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01566; NEUROMEDNU1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   402 AA;  45782 MW;  CB576083D9F7A095 CRC64;

  Query Match             50.8%;  Score 642;  DB 11;  Length 402;
  Best Local Similarity   54.1%;  Pred. No. 5.7e-52;
  Matches  125;  Conservative   36;  Mismatches   64;  Indels    6;  Gaps    2;

Qy      16 EDPFQKH-----LNSTEEYLAF-LCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQ 69
           :  |::|     || | | |     ||::   |||: | |: ||||| :|| | | |||:
Db       5 DSEFKEHFDLEDLNLTHEDLRLKYLGPQQVKQFLPICVTYLLIFVVGTLGNGLTCTVILR 64

Qy      70 HQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFA 129
           :|| ||||:|||||||||||||:|:|||:|||  ||||  |  ||||:  | |||| |
Db      65 QKAMHTPTNFYLFSLAVSDLLVLLVGLPLELYEMQHNYPFQLGAGGCYFRILLLETVCLA 124

Qy     130 SILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYF 189
           |:|::| :|||||||::|| :||   ||    |:|| :| |::|||||||:||:   |
Db     125 SVLNVTALSVERYVAVVHPLQAKSVMTRTHVRRMLGAIWVFAILFSLPNTSLHGLSPLYV 184

Qy     190 PNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           |    || | |||:::| : |  :|| |  ||: |||  ||||| |: ||:
Db     185 PCRGPVPDSVTCTLVRPQFFYKLVIQTTILLFFCLPMVTISVLYLLIGLRL 235


RESULT 8
Q9JJI5
ID   Q9JJI5      PRELIMINARY;      PRT;   412 AA.
AC   Q9JJI5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor FM-3.
GN   FM-3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20347213; PubMed=10783389;
RA   Fujii R., Hosoya M., Fukusumi S., Kawamata Y., Habata Y., Hinuma S.,
RA   Onda H., Nishimura O., Fujino M.;
RT   "Identification of neuromedin U as the cognate ligand of the orphan G
RT   protein-coupled receptor FM-3.";
RL   J. Biol. Chem. 275:21068-21074(2000).
DR   EMBL; AB038649; BAA99387.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01566; NEUROMEDNU1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   412 AA;  46784 MW;  2CAB56299E2EA7B8 CRC64;

  Query Match             50.8%;  Score 642;  DB 11;  Length 412;
  Best Local Similarity   54.1%;  Pred. No. 5.8e-52;
  Matches  125;  Conservative   36;  Mismatches   64;  Indels    6;  Gaps    2;

Qy      16 EDPFQKH-----LNSTEEYLAF-LCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQ 69
           :  |::|     || | | |     ||::   |||: | |: ||||| :|| | | |||:
Db      15 DSEFKEHFDLEDLNLTHEDLRLKYLGPQQVKQFLPICVTYLLIFVVGTLGNGLTCTVILR 74

Qy      70 HQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFA 129
           :|| ||||:|||||||||||||:|:|||:|||  ||||  |  ||||:  | |||| |
Db      75 QKAMHTPTNFYLFSLAVSDLLVLLVGLPLELYEMQHNYPFQLGAGGCYFRILLLETVCLA 134

Qy     130 SILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYF 189
```

```
          |:|::| :|||||||::|| :||   ||    |:|| :| |::|||||||:||:   |
Db    135 SVLNVTALSVERYVAVVHPLQAKSVMTRTHVRRMLGAIWVFAILFSLPNTSLHGLSPLYV 194

Qy    190 PNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
          |    || | |||:::| : |  :|| |  ||: |||  ||||| |: ||:
Db    195 PCRGPVPDSVTCTLVRPQFFYKLVIQTTILLFFCLPMVTISVLYLLIGLRL 245


RESULT 9
O55040
ID   O55040      PRELIMINARY;      PRT;   405 AA.
AC   O55040;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Orphan G protein-coupled receptor.
GN   GPR66.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McKee K.K., Tan C.T., Liu J., Palyha O.C., Feighner S.D.,
RA   Hreniuk D.L., Smith R.G., Howard A.D.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF044602; AAC02681.1; -.
DR   MGD; MGI:1341898; Gpr66.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01566; NEUROMEDNU1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   405 AA;  45609 MW;  F1BA493D3EB81F34 CRC64;

  Query Match             47.2%;  Score 596;  DB 11;  Length 405;
  Best Local Similarity   50.6%;  Pred. No. 1.1e-47;
  Matches  118;  Conservative   40;  Mismatches   67;  Indels    8;  Gaps    3;

Qy      9 WIYQQKLEDPFQKHLNSTEEYLAF-LCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVI 67
          | ||   ||     || |:| |     ||::   |:|: | |: ||||| :|| | | ||
Db     10 WPYQP--ED-----LNLTDEALRLKYLGPQQMKQFVPICVTYLLIFVVGTLGNGLTCTVI 62

Qy     68 LQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVC 127
          |::: |:||||:||||||||:||||:|:|||:||| :|||| |   |||:  | ||||
Db     63 LRNKTMRTPTNFYLFSLAVSDMLVLLVGLPLELYEMQQNYPFQLGASACYFRILLLETVC 122

Qy    128 FASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFH 187
           ||:|::| :|||||||:: | :||   ||    |::| :|  : |||||||:||:
Db    123 LASVLNVTALSVERYVAVVRPLQAKSVMTRAHVRRMVGAIWVLATLFSLPNTSLHGLSQL 182

Qy    188 YFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
            |    || || |::: ||  |  ::  |: ||: |||  ||||| |: ||:
Db    183 TVPCRGPVPDSAICSLVGPMDFYKLVVLTTALLFFCLPMVTISVLYLLIGLRL 235


RESULT 10
Q9VFN4
ID   Q9VFN4      PRELIMINARY;      PRT;   428 AA.
AC   Q9VFN4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG9918 protein.
GN   CG9918.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AE003703; AAF55016.1; -.
DR   FlyBase; FBgn0038201; CG9918.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   428 AA;  47198 MW;  9DD4A1E93046B6D7 CRC64;

  Query Match             33.1%;  Score 418;  DB 5;  Length 428;
  Best Local Similarity   40.9%;  Pred. No. 4.8e-31;
  Matches   85;  Conservative   49;  Mismatches   62;  Indels   12;  Gaps    4;

Qy     35 GPRRS--HFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVL 92
          || |      :||:|||  ||: ||:||:  |:|| ::::| | ||||||||||:|| |:|
Db     11 GPPRDPLAIVIPVTVVYSLIFITGVVGNISTCIVIKKNRSMHTATNYYLFSLAISDFLLL 70

Qy     93 LLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAK 152
          | |:| ||  :|  ||::||   |  :  | ||   |::|:||  :||||:|| |||  :
Db     71 LSGVPQEVSYIWSKYPYVFGEYICIGRGLLAETSANATVLTITAFTVERYIAICHPFLGQ 130

Qy    153 LQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNF 212
            |   ||:||: :||  ::: ::|  :  ||: ||        |   | ::: :  ::|
Db    131 AMSKLSRAIRIIVLVWIMAIVTAIPQAAQFGIE-HY-------SGVEQCGIVRVIVKHSF 182

Qy    213 IIQVTSFLFYLLPMTVISVLYYLMALRV 240
          |:::|:|:| ||::| ||| |: : :
```

```
Db    183 --QLSTFIFFLAPMSIILVLYLLIGVHL 208


RESULT 11
Q9VFW6
ID   Q9VFW6      PRELIMINARY;      PRT;   595 AA.
AC   Q9VFW6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG8795 protein.
GN   CG8795.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AE003699; AAF54929.1; -.
DR   FlyBase; FBgn0038139; CG8795.
DR   InterPro; IPR000923; BlueCu_1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PROSITE; PS00196; COPPER_BLUE; UNKNOWN_1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   595 AA;  65091 MW;  3E8E183BADF5471E CRC64;
```

```
  Query Match           27.9%;  Score 352;  DB 5;  Length 595;
  Best Local Similarity 37.9%;  Pred. No. 9.5e-25;
  Matches   75;  Conservative   43;  Mismatches   66;  Indels   14;  Gaps    5;

Qy      45 VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMW 104
           :|| |  ||: ||:||:: |:|| ::  | | ||:|||:||:||:::|  ||| ::| :|
Db      65 LSVGYALIFIAGVLGNLITCIVISRNNFMHTATNFYLFNLAISDMILLCSGMPQDLYNLW 124

Qy     105 R--NYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALR 162
             |||  |    |  :: | ||   |::|:||  :||||:|| ||||    |   ||::
Db     125 HPDNYP--FSDSICILESVLSETAANATVLTITAFTVERYIAICHPFRQHTMSKLSRAVK 182

Qy     163 ILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFY 222
           :  :|  ::| :||       |:|      |       :||:    : :  :  |: |||:
Db     183 FIFAIWIAALLLALP----QAIQFSVVMQGM----GTSCTMKNDFFAH--VFAVSGFLFF 232

Qy     223 LLPMTVISVLYYLMALRV 240
           ||| | ||| |: :::
Db     233 GGPMTAICVLYVLIGVKL 250


RESULT 12
Q9VFW5
ID   Q9VFW5      PRELIMINARY;      PRT;   660 AA.
AC   Q9VFW5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG8784 protein.
GN   CG8784.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
```

```
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AE003699; AAF54930.1; -.
DR   FlyBase; FBgn0038140; CG8784.
DR   InterPro; IPR000923; BlueCu_1.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PROSITE; PS00196; COPPER_BLUE; UNKNOWN_1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   660 AA;  72277 MW;  C52D1EA7E4ACD439 CRC64;

  Query Match             26.7%;  Score 337;  DB 5;  Length 660;
  Best Local Similarity   34.1%;  Pred. No. 2.6e-23;
  Matches   84;  Conservative   42;  Mismatches   86;  Indels   34;  Gaps    7;

Qy      17 DPFQKH----LNSTEEYLAFLCG-------------PRRSHFFL--PVSVVYVPIFVVG 56
           |  |  |    || | | |: | |             |      |   ::| |  ||| |
Db      62 DKFLTHVAHLLNITTENLSNLLGSTNGTNASTMAADSPVDESLTLRTALTVCYALIFVAG 121

Qy      57 VIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYP--FLFGPV 114
           |:||:: |:|| ::  | | ||:|||:||||||::|: |:| |:| :|  ||  : |
Db     122 VLGNLITCIVISRNNFMHTATNFYLFNLAVSDLILLVSGIPQELYNLW--YPDMYPFTDA 179

Qy     115 GCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLF 174
           |   : | |    |::|:||  :||||:|| ||||    |   ||:: :  :|  : |
Db     180 MCIMGSVLSEMAANATVLTITAFTVERYIAICHPFRQHTMSKLSRAIKFIFAIWLAAFLL 239

Qy     175 SLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYY 234
           :||      : :           :||:      |  :  |: |:|:  ||| | |||
Db     240 ALPQAMQFSVVYQ--------NEGYSCTMEND--FYAHVFAVSGFIFFGGPMTAICVLYV 289

Qy     235 LMALRV 240
           |: :::
Db     290 LIGVKL 295


RESULT 13
Q18701
ID   Q18701      PRELIMINARY;      PRT;   378 AA.
AC   Q18701;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   C48C5.1 protein.
GN   C48C5.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Favello A.;
RT   "The sequence of C. elegans cosmid C48C5.";
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U39994; AAB37017.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   378 AA;  43631 MW;  458DC60CCBF88F86 CRC64;

  Query Match             26.3%;  Score 332.5;  DB 5;  Length 378;
  Best Local Similarity   33.6%;  Pred. No. 4.1e-23;
  Matches   87;  Conservative   37;  Mismatches   98;  Indels   37;  Gaps    6;

Qy      20 QKHLNSTEEY---LAFLCGPRRSH----------------------FFLPVSVVYVPIF 53
           |  ||:||:    ||| |    ||                        |: :|: ||
Db       3 QACLNTTEDQCDCLAFNCPIVYSHSESEKEACYMEHCFISKRALDDVTLYKVTALYIFIF 62

Qy      54 VVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGP 113
           :||||||   |||: :|  |||  : || :||||||: | :|:| ||   |  ||: |
Db      63 LVGVIGNTTTCLVMKKHPMMKTHASMYLMNLAVSDLVTLCVGLPFEVMMNWNQYPWPFPD 122

Qy     114 VGCYFKTALFETVCFASILSITTVSVERYVAILHP-FRAKLQSTRRRALRILGIVWGFSV 172
           |  |  : ||    |||:|   ::|||||: || |  |:|  :|    |:|  | ||:
Db     123 YICNLKALIAETTSSVSILTILIFAIERYVAVCHPLFLMKVQPFKRNIGTIIGFTWIFSI 182

Qy     173 LFSLPNTSIHGIKF--HYFP---NGSLVPGSATCTVI-----KPMWIYNFIIQVTSFLFY 222
           | ::|    |   :    :|   |   |  |  | :      |    :  :   ::  |:
Db     183 LCAMPFAIHHRADYIMKSWPGTDNRIPVKSSKMCMIAVMFEPKLASTFKILFHFSAIAFF 242

Qy     223 LLPMTVISVLYYLMALRVS 241
           ||:  | :||  :| :||
Db     243 ALPLFTIVILYARIACKVS 261


RESULT 14
O17239
ID   O17239      PRELIMINARY;      PRT;   418 AA.
AC   O17239;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 47.7 kDa protein.
GN   K10B4.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Sammons L., Wohldmann P., Antoniou B.;
RT   "The sequence of C. elegans cosmid K10B4.";
```

```
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF025463; AAB71009.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   418 AA;  47735 MW;  CE0416539CA3BB27 CRC64;

  Query Match             24.0%;  Score 303.5;  DB 5;  Length 418;
  Best Local Similarity   30.2%;  Pred. No. 2.3e-20;
  Matches   76;  Conservative   54;  Mismatches   89;  Indels   33;  Gaps    6;

Qy      22 HLNSTEEYLAFLCGPR--RSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNY 79
           :::   ||:    | |   :   :|  ::|  ||::|: ||:  |:||  :::|  ||||
Db       9 NVSEITEYVLSTLGERCQSAGIVIPTVIIYGTIFLLGLFGNICTCIVIAANKSMHNPTNY 68

Qy      80 YLFSLAVSDLLVLLLGMPLEVYE-MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVS 138
           |||||||||:: |:||:|:| |: :  :||: |    |  :  | |   :|||: |   |
Db      69 YLFSLAVSDIIALILGLPMEFYQSLDYSYPYRFSEGICKARAFLIEFTSYASIMIICCFS 128

Qy     139 VERYVAILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGI-------------- 184
           ||::|| || |:|: ||  ||  :: : |  | : :||   |  |
Db     129 FERWLAICHPLRSKIFSTLWRANVLIILAWTISFVCALPIAFIVQINKLPLPEDAKYQPW 188

Qy     185 --KFHYFP----NGSLVPGSAT---------CTVIKPMWIYNFIIQVTSF-LFYLLPMTV 228
             |  :|     |  : | |           | : :       :| : :| :|:::|
Db     189 TNKVSFFAVGVLNNRIFPVSTDGIFVLHTEFCAMNQSRPDQQKMIIIFAFTVFFVIPAIA 248

Qy     229 ISVLYYLMALRV 240
           | ::|  :|:::
Db     249 IVIMYAHIAVQL 260


RESULT 15
O93412
ID   O93412      PRELIMINARY;      PRT;   374 AA.
AC   O93412;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Orphan G protein-coupled receptor.
OS   Sphoeroides nephelus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
     Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
     Tetraodontidae; Sphoeroides.
OX   NCBI_TaxID=39110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20092336; PubMed=10628755;
RA   Palyha O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA   Gao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
RA   Patchett A.A., Howard A.D., Smith R.G.;
RT   "Ligand activation domain of human orphan growth hormone (GH)
RT   secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL   Mol. Endocrinol. 14:160-169(2000).
DR   EMBL; AF082209; AAC33472.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR000524; HTH_GntR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01565; NEUROMEDINUR.
DR   PRINTS; PR01566; NEUROMEDNU1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
KW   Receptor.
SQ   SEQUENCE   374 AA;  42324 MW;  2CF9304F004C7A16 CRC64;

  Query Match             24.0%;  Score 303;  DB 13;  Length 374;
  Best Local Similarity   32.2%;  Pred. No. 2.3e-20;
  Matches   74;  Conservative   51;  Mismatches   71;  Indels   34;  Gaps    6;

Qy      36 PRRSHFFLP----VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLV 91
           |  ::: :|    ::|    :| |||:|||:  ||: ::: |:| || || |:|||| :
Db      26 PPLNYYSIPLLAVITVACTVLFTVGVVGNVMTILVVSRYRDMRTTTNLYLCSMAVSDLFI 85

Qy      92 LLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRA 151
            :  |||::| |||  |: ||   |     : |:  :::|| || :|||||:||  | ||
Db      86 FVC-MPLDLYRMWRYRPWRFGDALCKLFQFVSESCTYSTILCITALSVERYLAICFPLRA 144

Qy     152 KLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFH--YFPNGSLVPGSA---------T 200
           |   |:||   :: ::|  |:| : |   : |::     ||| | :  |:
Db     145 KALVTKRRVRALILLLWTVSLLSAGPVFVMVGVEKDSIMFPNSSDLNESSWPLEAVDTRE 204

Qy     201 CTVIK----------PMWIYNFIIQVTSFLFYLLPMTVISVLYYLMALRV 240
           | : :           :|:        | :|: :|:  ::||| |:  |:
Db     205 CRMTQYAVESGLMEAMVWL--------SSVFFFMPVFCLTVLYGLIGRRL 246


Search completed: January 17, 2003, 05:06:39
Job time : 97 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:19:20 ; Search time 265 Seconds
(without alignments)
6195.120 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc.....cactcagagtgagtatctag 729

[redacted]ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	729	100.0	729	22	AAF85107	Nucleotide sequenc
c 2	729	100.0	801	22	AAH50977	Human nGPCR15 codi
c 3	729	100.0	813	24	AAS98102	Human DNA for pote
4	717.8	98.5	1239	24	AAS98055	Human DNA for pote
5	717.8	98.5	1245	22	AAH43072	Nucleotide sequenc
6	717.8	98.5	1245	22	AAH43075	Nucleotide sequenc
7	717.8	98.5	1248	21	AAD01123	Human orphan G pro
8	717.8	98.5	1248	21	AAA46022	Human G protein co
9	717.8	98.5	1248	22	AAH49526	Human GTP-binding
10	717.8	98.5	1248	24	ABK47909	Human G-protein co
11	717.8	98.5	1298	22	AAD08008	Human G-protein co
12	717.8	98.5	1344	24	AAD23701	Human NMUR2 cDNA.
13	717.8	98.5	1360	21	AAZ33297	Human neurotensin-
14	717.8	98.5	1594	22	AAF80326	Splice variant of
15	717.8	98.5	1594	22	AAF80327	Splice variant of
16	717.8	98.5	1658	22	AAF80322	Nucleotide sequenc
17	717.8	98.5	1658	22	AAF80323	Nucleotide sequenc
18	717.8	98.5	1658	22	AAF80324	Nucleotide sequenc
19	717.8	98.5	1658	22	AAF80325	Nucleotide sequenc
20	471.4	64.7	1185	22	AAH43078	Nucleotide sequenc
21	471.4	64.7	1188	24	AAD23702	Rat NMUR2 cDNA. R
22	469.8	64.4	1231	22	AAD08024	Rat G-protein coup
23	282.4	38.7	1349	22	AAS59858	Human novel cytoki
24	282.4	38.7	1535	22	AAS59815	Human novel cytoki
25	280.8	38.5	1209	22	AAH45072	Human FM-3 coding
26	280.8	38.5	1209	22	AAF76231	Human G-protein co
27	280.8	38.5	1212	21	AAA30663	Human G protein-co
28	280.8	38.5	1212	21	AAA30739	DNA encoding human
29	280.8	38.5	1318	22	AAD08007	Human G-protein co
30	279.2	38.3	1212	21	AAZ49707	Human growth hormo
31	277.6	38.1	2374	24	AAD24956	Human G-protein co
32	255	35.0	1239	22	AAH45076	Rat FM-3 coding se
33	255	35.0	1292	22	AAD08025	Rat G-protein coup
34	255	35.0	1371	22	AAD08026	Rat G-protein coup
35	243	33.3	1215	22	AAH45073	Murine FM-3 coding
36	243	33.3	1526	21	AAZ49706	Mouse growth hormo
37	135.8	18.6	1203	22	AAF85448	Nucleotide sequenc
38	134.8	18.5	1161	22	AAF83684	Short form of moti
39	134.8	18.5	1239	21	AAZ45403	cDNA encoding the
40	134.8	18.5	1239	22	AAF83683	Long form of motil
41	134.8	18.5	1239	22	AAF85449	Nucleotide sequenc
42	134.8	18.5	1390	21	AAZ45404	cDNA encoding the
43	134.8	18.5	1506	22	AAI66989	Human GPR38 varian
44	134.8	18.5	2040	21	AAA46116	Human G protein co
45	134.8	18.5	3066	21	AAZ45402	Genomic sequence o

ALIGNMENTS

RESULT 1
AAF85107
ID AAF85107 standard; DNA; 729 BP.
XX
AC AAF85107;
XX
DT 09-JUL-2001 (first entry)
XX
DE Nucleotide sequence of human g-protein coupled receptor PFI-002.
XX
KW G-protein coupled receptor; obesity; signal transduction; diabetes;
KW metabolic disease; neurological disease; psychotherapy; dermatology;
KW urogenital disease; inflammation; cancer; tissue repair; photoageing;
KW skin pigmentation; frailty; osteoporosis; cardiovascular disease;
KW gastrointestinal disease; infection; allergy; respiratory disease;
KW sensory organ disorder; sleep disorder; hair loss; gene therapy;
KW PFI-002; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..729
FT /*tag= a
FT /product= "g-protein coupled receptor"
XX
PN EP1090990-A1.
XX
PD 11-APR-2001.
XX
PF 06-OCT-2000; 2000EP-0308852.
XX
PR 08-OCT-1999; 99GB-0023888.

XX
PA (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
PI Harland L;
XX
DR WPI; 2001-302046/32.
DR P-PSDB; AAB68333.
XX
PT New human G-protein coupled receptor (GPCR) polynucleotides and
PT polypeptides, for screening modulators of the polypeptide useful in
PT treating diseases associated with signal transduction, e.g. cancer,
PT inflammation, or especially, obesity -
XX
PS Claim 1; Page 42; 53pp; English.
XX
CC The present sequence encodes a human G-protein coupled receptor. The
CC G-protein coupled receptor polynucleotide and polypeptide are useful
CC as pharmaceuticals or in the manufacture of medicaments for the
CC treatment of obesity. They are useful in the diagnosis and treatment
CC of diseases and disorders associated with signal transduction such as
CC obesity, diabetes and metabolic disease, neurological disease,
CC psychotherapeutics, urogenital disease, inflammation, cancer, tissue
CC repair, dermatology, skin pigmentation, photoageing, frailty,
CC osteoporosis, cardiovascular disease, gastrointestinal disease,
CC infection, allergy and respiratory disease, sensory organ disorders,
CC sleep disorders and hair loss. The polynucleotide may also by useful
CC in gene therapy.
XX
SQ Sequence 729 BP; 135 A; 245 C; 167 G; 182 T; 0 other;

Query Match 100.0%; Score 729; DB 22; Length 729;
Best Local Similarity 100.0%; Pred. No. 4e-148;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720

Qy    721 AGTATCTAG 729
          |||||||||
Db    721 AGTATCTAG 729
```

RESULT 2
AAH50977/c
ID AAH50977 standard; DNA; 801 BP.
XX
AC AAH50977;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPCR15 coding sequence.
XX
KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; ds.
XX
OS Homo sapiens.
XX
PN WO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31581.
XX
PR 16-NOV-1999; 99US-0165838.
PR 17-NOV-1999; 99US-0166071.
PR 19-NOV-1999; 99US-0166678.
PR 28-DEC-1999; 99US-0173396.
PR 22-FEB-2000; 2000US-0184129.
PR 28-FEB-2000; 2000US-0185421.
PR 28-FEB-2000; 2000US-0185554.
PR 02-MAR-2000; 2000US-0186530.
PR 03-MAR-2000; 2000US-0186811.
PR 09-MAR-2000; 2000US-0188114.
PR 17-MAR-2000; 2000US-0190310.
PR 21-MAR-2000; 2000US-0190800.
PR 20-APR-2000; 2000US-0198568.
PR 02-MAY-2000; 2000US-0201190.
PR 08-MAY-2000; 2000US-0203111.
PR 25-MAY-2000; 2000US-0207094.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
XX
DR WPI; 2001-389826/41.
DR P-PSDB; AAG80937.
XX
PT New G protein-coupled receptor (nGPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
PS Claim 4; Page 78-79; 261pp; English.
XX

```
CC   The present invention relates to novel G protein-coupled receptors
CC   (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC   28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC   sequence is the coding sequence for one such G protein-coupled receptor.
CC   GPCRs are also known as seven transmembrane receptors and function in
CC   signal transduction. The nGPCRx coding sequences are useful for
CC   screening a human to diagnose a disorder affecting the brain or a genetic
CC   predisposition, specifically schizophrenia. nGPCRx are useful for
CC   identifying compounds useful for treating schizophrenia. Detection of
CC   nGPCRx in a sample is useful as a diagnostic tool for diseases or
CC   disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC   CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC   diseases, proliferative disorders and hormonal disorders. Modulators of
CC   nGPCRx activity have the utility for treating neurological disorders,
     including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
     disorder/attention deficit disorder), and neuronal disorders such as
     Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC   Additional disorders include inflammatory conditions (e.g. Crohn's
CC   disease), rheumatoid arthritis, autoimmune disorders, cancers,
CC   respiratory ailments such as asthma, and inflammatory diseases e.g.
CC   inflammatory bowel disease.
XX
SQ   Sequence 801 BP; 200 A; 187 C; 261 G; 153 T; 0 other;

  Query Match            100.0%;  Score 729;  DB 22;  Length 801;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-148;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 736

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 676

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 616

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 556

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 496

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 436

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 376

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 316

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 256

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 196

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 136

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 76

Qy    721 AGTATCTAG 729
          |||||||||
Db     75 AGTATCTAG 67


RESULT 3
AAS98102/c
ID   AAS98102 standard; DNA; 813 BP.
XX
AC   AAS98102;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human DNA for potential G protein-coupled receptor #59.
XX
KW   Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
KW   Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
KW   atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
KW   chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
KW   depression; epilepsy; macular degeneration; lymphoma; melanoma;
KW   multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KW   psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;
KW   tuberculosis; cognition disorder; memory disorder; anorexia;
KW   hormonal release disorder; cardiovascular activity disorder;
KW   pain perception disorder; obesity; diabetes; obesity;
KW   diabetes; hyperlipidaemia; stroke; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200185791-A1.
XX
PD   15-NOV-2001.
XX
PF   11-MAY-2001; 2001WO-US15332.
XX
PR   11-MAY-2000; 2000US-203217P.
PR   18-MAY-2000; 2000US-205945P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Brown JP,  Miller M,  Burmer G,  Fabre-Suver C,  Pritchard D;
XX
DR   WPI; 2002-066595/09.
XX
PT   Novel G protein-coupled receptor polypeptides including galanin
PT   receptor polypeptides useful for identifying modulators that are useful
PT   for treating Alzheimer's disease, psoriasis, melanoma, multiple
PT   sclerosis, stroke  -
XX
PS   Claim 2; Page 79; 144pp; English.
XX
CC   The invention relates to an isolated polypeptide encoded by a
CC   nucleic acid molecule that is at least 80% identical to the G
CC   protein-coupled (GPCR) polynucleotides included in the specification.
CC   Also included are probes based on the GPCR sequences (including
CC   antisense probes), a host cell comprising an expression vector comprising
CC   the GPCR sequence, antibodies raised against the polypeptides,
CC   and methods of identifying modulators of the polypeptides. The
CC   polypeptides are useful for identifying modulator compounds which
CC   function as modulators, activators, repressors, agonists or antagonists
CC   of the novel GPCR polypeptides including the GAL4 polypeptide. The
CC   antibodies and nucleic acid probes as described above can be used to
CC   detect the presence of the polypeptides and nucleic acids and are used to
CC   diagnose a variety of diseases or disorders in which GPCRs are involved
CC   e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,
CC   atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,
CC   chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,
CC   depression, epilepsy, macular degeneration, lymphoma, melanoma,
CC   multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,
CC   psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,
```

```
CC   tuberculosis and many other diseases listed in the specification. The
CC   probes and antibodies are also useful for diagnosing cognition and memory
CC   disorders, anorexia, hormonal release disorders, cardiovascular activity
CC   disorders, pain perception disorders, obesity, diabetes, Alzheimer's
CC   disease. Preferably, compounds that decrease or increase
CC   the expression of galanin receptor (GAL4) can be used to treat obesity,
CC   diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is
CC   useful for treating the  above mentioned disorders by gene therapy
CC   techniques. The present sequence is a novel GPCR polynucleotide of the
CC   invention.
XX
SQ   Sequence 813 BP; 206 A; 190 C; 261 G; 156 T; 0 other;

  Query Match             100.0%;  Score 729;  DB 24;  Length 813;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-148;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 736

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 676

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 616

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 556

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 496

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 436

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 376

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 316

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 256

      541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 196

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 136

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 76

Qy    721 AGTATCTAG 729
          |||||||||
Db     75 AGTATCTAG 67


RESULT 4
AAS98055
ID   AAS98055 standard; DNA; 1239 BP.
XX
AC   AAS98055;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Human DNA for potential G protein-coupled receptor #13.
XX
KW   Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
KW   Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
KW   atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
KW   chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
KW   depression; epilepsy; macular degeneration; lymphoma; melanoma;
KW   multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KW   psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;
KW   tuberculosis; cognition disorder; memory disorder; anorexia;
KW   hormonal release disorder; cardiovascular activity disorder;
KW   pain perception disorder; obesity; diabetes; obesity;
KW   diabetes; hyperlipidaemia; stroke; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200185791-A1.
XX
PD   15-NOV-2001.
XX
PF   11-MAY-2001; 2001WO-US15332.
XX
PR   11-MAY-2000; 2000US-203217P.
PR   18-MAY-2000; 2000US-205945P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Brown JP,  Miller M,  Burmer G,  Fabre-Suver C,  Pritchard D;
XX
DR   WPI; 2002-066595/09.
XX
PT   Novel G protein-coupled receptor polypeptides including galanin
PT   receptor polypeptides useful for identifying modulators that are useful
PT   for treating Alzheimer's disease, psoriasis, melanoma, multiple
PT   sclerosis, stroke  -
XX
PS   Disclosure; Page 120-121; 144pp; English.
XX
CC   The invention relates to an isolated polypeptide encoded by a
CC   nucleic acid molecule that is at least 80% identical to the G
CC   protein-coupled (GPCR) polynucleotides included in the specification.
CC   Also included are probes based on the GPCR sequences (including
CC   antisense probes), a host cell comprising an expression vector comprising
CC   the GPCR sequence, antibodies raised against the polypeptides,
CC   and methods of identifying modulators of the polypeptides. The
CC   polypeptides are useful for identifying modulator compounds which
CC   function as modulators, activators, repressors, agonists or antagonists
CC   of the novel GPCR polypeptides including the GAL4 polypeptide. The
CC   antibodies and nucleic acid probes as described above can be used to
CC   detect the presence of the polypeptides and nucleic acids and are used to
CC   diagnose a variety of diseases or disorders in which GPCRs are involved
CC   e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,
CC   atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,
CC   chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,
CC   depression, epilepsy, macular degeneration, lymphoma, melanoma,
CC   multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,
CC   psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,
CC   tuberculosis and many other diseases listed in the specification. The
CC   probes and antibodies are also useful for diagnosing cognition and memory
CC   disorders, anorexia, hormonal release disorders, cardiovascular activity
CC   disorders, pain perception disorders, obesity, diabetes, Alzheimer's
CC   disease. Preferably, compounds that decrease or increase
CC   the expression of galanin receptor (GAL4) can be used to treat obesity,
CC   diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is
CC   useful for treating the  above mentioned disorders by gene therapy
CC   techniques. The present sequence is a novel GPCR polynucleotide of the
CC   invention.
XX
```

```
SQ   Sequence 1239 BP; 267 A; 392 C; 265 G; 315 T; 0 other;

  Query Match            98.5%;  Score 717.8;  DB 24;  Length 1239;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120

       121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 720

Qy    721 A 721
          |
Db    721 A 721


RESULT 5
AAH43072
ID   AAH43072 standard; DNA; 1245 BP.
XX
AC   AAH43072;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of a human TGR-1 protein.
XX
KW   TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1245
FT                   /*tag= a
FT                   /product= "TGR-1"
FT                   /note= "no termination codon given"
XX
PN   WO200157524-A1.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-JP00746.
XX
PR   04-FEB-2000; 2000JP-0032773.
PR   24-FEB-2000; 2000JP-0052252.
PR   30-MAR-2000; 2000JP-0097896.
PR   19-JUN-2000; 2000JP-0187536.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Hinuma S,  Shintani Y,  Hosoya M,  Fujii R,  Moriya T,  Matsui H;
PI   Okubo S;
XX
DR   WPI; 2001-488917/53.
DR   P-PSDB; AAG63353.
XX
PT   Identifying predicted or actual structures of two or more members of a
PT   chemical or physical library by mass spectrometry comprising
PT   correlating molecular mass measurements of two or more members with a
PT   shared chemical history -
XX
PS   Claim 9; Page 79-80; 95pp; Japanese.
XX
CC   The present sequence encodes a human TGR-1 protein. The specification
CC   describes a method of screening a compound, which is capable of binding
CC   properties of neuromedin U to TGR-1. The method is useful for screening
CC   preventatives and remedies for hypertension, stress diseases, etc..
CC   TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ   Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;

  Query Match            98.5%;  Score 717.8;  DB 22;  Length 1245;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy    721 A 721
          |
Db    730 A 730
```

RESULT 6
AAH43075
ID AAH43075 standard; DNA; 1245 BP.
XX
AC AAH43075;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1245
FT /*tag= a
FT /product= "TGR-1"
FT /note= "no termination codon given"
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-JP00746.
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
30-MAR-2000; 2000JP-0097896.
19-JUN-2000; 2000JP-0187536.

PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI; 2001-488917/53.
DR P-PSDB; AAG63366.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
PS Disclosure; Page 85-87; 95pp; Japanese.
XX
CC The present sequence encodes a human TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventatives and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ Sequence 1245 BP; 268 A; 394 C; 268 G; 315 T; 0 other;

Query Match 98.5%; Score 717.8; DB 22; Length 1245;
Best Local Similarity 99.7%; Pred. No. 1.2e-145;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy    721 A 721
          |
Db    730 A 730
```

RESULT 7
AAD01123
ID AAD01123 standard; cDNA; 1248 BP.
XX
AC AAD01123;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human orphan G protein-coupled receptor hRUP6 cDNA.
XX
KW Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;

```
KW   transmembrane receptor; signal cascade; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1248
FT                   /*tag=  a
FT                   /product= "hRUP6"
FT                   /note= "Human orphan G protein-coupled receptor"
XX
PN   WO200031258-A2.
XX
PD   02-JUN-2000.
XX
     13-OCT-1999;   99WO-US23687.

     20-NOV-1998;   98US-0109213.
PR   16-FEB-1999;   99US-0120416.
PR   26-FEB-1999;   99US-0121852.
PR   12-MAR-1999;   99US-0123946.
PR   12-MAR-1999;   99US-0123949.
PR   28-MAY-1999;   99US-0136436.
PR   28-MAY-1999;   99US-0136437.
PR   28-MAY-1999;   99US-0136439.
PR   28-MAY-1999;   99US-0136567.
PR   28-MAY-1999;   99US-0137127.
PR   28-MAY-1999;   99US-0137131.
PR   29-JUN-1999;   99US-0141448.
PR   29-SEP-1999;   99US-0156555.
PR   29-SEP-1999;   99US-0156633.
PR   29-SEP-1999;   99US-0156634.
PR   29-SEP-1999;   99US-0156653.
PR   01-OCT-1999;   99US-0157280.
PR   01-OCT-1999;   99US-0157281.
PR   01-OCT-1999;   99US-0157282.
PR   01-OCT-1999;   99US-0157293.
PR   01-OCT-1999;   99US-0157294.
PR   12-OCT-1999;   99US-0416760.
PR   12-OCT-1999;   99US-0417044.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Liaw CW,  Lin I;

     WPI; 2000-400068/34.
DR   P-PSDB; AAY71296.
XX
PT   Novel human orphan G protein-coupled receptors and the encoding cDNAs
PT   for use in the identification of G protein-coupled receptor agonists -
XX
PS   Claim 21; Page 56-57; 102pp; English.
XX
CC   The present sequence is a cDNA encoding hRUP6, an endogenous human
CC   orphan G protein-coupled receptor (GPCR). The full length hRUP6 cDNA
CC   was cloned by RT-PCR using human thymus cDNA as template.
CC   The orphan GPCR of the invention, like all GPCRs has seven transmembrane
CC   alpha helices with an extracellular N-terminus and an intracellular
CC   C-terminus. However, no endogenous ligands has yet been identified for
CC   the proteins of the invention. The orphan GPCRs may be used in the
CC   identification of their endogenous ligands, and to screen potential GPCR
CC   agonists and antagonists for use as pharmaceutical agents. The proteins
CC   may also be used in the study of GPCR-mediated signalling cascades, and
CC   to elucidate their precise role in normal and diseased human conditions.
CC   Nucleic acid encoding human orphan GPCRs may be used for tissue
CC   localisation expression analysis to provide information about their
CC   function in healthy and pathological states.
XX
SQ   Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

  Query Match            98.5%;  Score 717.8;  DB 21;  Length 1248;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy     721 A 721
           |
Db     730 A 730


RESULT 8
AAA46022
ID   AAA46022 standard; cDNA; 1248 BP.
XX
AC   AAA46022;
XX
DT   22-AUG-2000  (first entry)
XX
DE   Human G protein coupled receptor hRUP6 encoding cDNA SEQ ID NO:11.
XX
KW   Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW   identification; agonist; screening; therapeutic; pharmaceutical;
KW   mutant; ss.
XX
OS   Homo sapiens.
XX
PN   WO200022131-A2.
XX
PD   20-APR-2000.
XX
```

```
PF   13-OCT-1999;   99WO-US24065.
XX
PR   13-OCT-1998;   98US-0170496.
PR   12-NOV-1998;   98US-0108029.
PR   20-NOV-1998;   98US-0109213.
PR   27-NOV-1998;   98US-0110060.
PR   16-FEB-1999;   99US-0120416.
PR   26-FEB-1999;   99US-0121852.
PR   12-MAR-1999;   99US-0123944.
PR   12-MAR-1999;   99US-0123945.
PR   12-MAR-1999;   99US-0123946.
PR   12-MAR-1999;   99US-0123948.
PR   12-MAR-1999;   99US-0123949.
PR   12-MAR-1999;   99US-0123951.
PR   28-MAY-1999;   99US-0136436.
PR   28-MAY-1999;   99US-0136437.
PR   28-MAY-1999;   99US-0136439.
PR   28-MAY-1999;   99US-0137127.
PR   28-MAY-1999;   99US-0137131.
PR   28-MAY-1999;   99US-0137567.
PR   30-JUN-1999;   99US-0141448.
PR   27-AUG-1999;   99US-0151114.
PR   03-SEP-1999;   99US-0152524.
PR   29-SEP-1999;   99US-0156633.
PR   29-SEP-1999;   99US-0156555.
PR   29-SEP-1999;   99US-0156634.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Behan DP,  Lehmann-Bruinsma K,  Chalmers DT,  Chen R,  Dang HT;
PI   Gore M,  Liaw CW,  Lin I,  Lowitz K,  White C;
XX
DR   WPI; 2000-317986/27.
DR   P-PSDB; AAB02830.
XX
PT   Non-endogenous, human G protein-coupled receptors for screening
PT   receptor, inverse or partial agonists useful as therapeutic agents  -
XX
PS   Example 1; Page 85-86; 187pp; English.
XX
CC   The present invention describes transmembrane receptors, preferably
CC   human G protein coupled receptors (GPCR), for which the endogenous
CC   ligand is unknown (orphan GPCR receptors). More specifically the present
CC   invention relates to non-endogenous, constitutively activated versions
CC   of a human GPCR. These non-endogenous human GPCRs can be useful for
CC   the direct identification of candidate compounds as receptors agonists,
CC   inverse agonists or partial agonists for use as pharmaceutical agents.
CC   AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

 Query Match             98.5%;  Score 717.8;  DB 21;  Length 1248;
 Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
 Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy     721 A 721
           |
Db     730 A 730


RESULT 9
AAH49526
ID   AAH49526 standard; DNA; 1248 BP.
XX
AC   AAH49526;
XX
DT   21-SEP-2001  (first entry)
XX
DE   Human GTP-binding protein-coupled receptor GPRv39 coding sequence.
XX
KW   GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
KW   muscular; urinary; circulatory; anorectic; gene therapy; human;
KW   guanosine triphosphate; G-protein; ds.
XX
OS   Homo sapiens.
XX
PN   WO200148189-A1.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-JP09409.
XX
PR   28-DEC-1999;   99JP-0375152.
PR   31-MAR-2000; 2000JP-0101339.
PR   23-MAY-2000; 2000JP-0155978.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Matsumoto S,  Oda T,  Saito Y,  Morikawa N,  Yoshida K,  Suwa M;
PI   Sugiyama T;
XX
DR   WPI; 2001-425663/45.
DR   P-PSDB; AAG64297.
XX
PT   Family of guanosine triphosphate binding protein coupled receptors and
PT   genes encoding them for treatment and prevention of diseases associated
PT   with these receptors  -
```

```
XX
PS   Claim 1; Page 105; 137pp; Japanese.
XX
CC   The present sequence is the coding sequence for a human guanosine
CC   triphosphate (GTP)-binding protein-coupled receptor. The receptor is
CC   useful for the investigation, diagnosis, treatment and prevention of
CC   diseases associated with GTP-binding protein-coupled receptors, including
CC   neurological, circulatory, digestive system, immune system, muscle and
CC   urinary system disorders. GTP-binding proteins are also known as
CC   G-proteins.
XX
SQ   Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

  Query Match             98.5%;  Score 717.8;  DB 22;  Length 1248;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy     721 A 721
          |
Db     730 A 730


RESULT 10
ABK47909
ID   ABK47909 standard; cDNA; 1248 BP.
XX
AC   ABK47909;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human G-protein coupled receptor AXOR34 cDNA.
XX
KW   Human; G protein-coupled; AXOR34; gene; ss; bacterial infection; cancer;
KW   fungal infection; protozoan infection; viral infection; hypertension;
KW   pain; diabetes; obesity; anorexia; bulimia; Parkinson's disease; stroke;
KW   acute heart failure; hypotension; urinary retention; osteoporosis; ulcer;
KW   angina pectoris; myocardial infarction; benign prostatic hypertrophy;
KW   allergy; migraine; vomiting; psychotic disorder; neurological disorder;
KW   dyskinesia; receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1248
FT                   /*tag=  a
FT                   /product= "Human AXOR34"
XX
PN   GB2363793-A.
XX
PD   09-JAN-2002.
XX
PF   10-APR-2001; 2001GB-0008983.
XX
PR   10-APR-2000; 2000US-0545944.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Ames RS,  Sarau H,  Vawter L,  Elshourbagy N,  Shabon U;
PI   Michalovich D;
XX
DR   WPI; 2002-306791/35.
DR   P-PSDB; AAU77155.
XX
PT   New G protein-coupled receptor AXOR34, useful for diagnosis and
PT   treatment of e.g. infections and cancer -
XX
PS   Claim 1; Fig 1; 46pp; English.
XX
CC   The invention relates to a human G protein-coupled receptor AXOR34 and
CC   the polynucleotide encoding it. The sequences of the invention are used
CC   for treating and/or a variety of diseases, including bacterial, fungal,
CC   protozoan and viral infections (particularly infections caused by HIV),
CC   pain, cancers, diabetes, obesity, anorexia, bulimia, Parkinson's disease,
CC   acute heart failure, hypotension, hypertension, urinary retention,
CC   osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,
CC   allergies, benign prostatic hypertrophy, migraine, vomiting, psychotic
CC   and neurological disorders (including anxiety, schizophrenia, manic
CC   depression, depression, delirium, dementia and severe mental retardation)
CC   and dyskinesias (such as Huntington's disease and Gilles de la Tourette's
CC   syndrome). This sequence represents cDNA encoding the human AXOR34
CC   polypeptide.
XX
SQ   Sequence 1248 BP; 269 A; 392 C; 269 G; 318 T; 0 other;

  Query Match             98.5%;  Score 717.8;  DB 24;  Length 1248;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy     721 A 721
           |
Db     730 A 730
```

RESULT 11
AAD08008
ID AAD08008 standard; cDNA; 1298 BP.
XX
AC AAD08008;
XX
DT 07-AUG-2001 (first entry)
XX
DE Human G-protein coupled receptor, SNORF72 cDNA.
XX
KW Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
KW NMU; inflammation; arthritis; autoimmune disease; septicaemia; psychotic;
KW mental retardation; transplant rejection; neurological disorder; anxiety;
KW respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
KW gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
KW ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
KW dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
KW Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
KW cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
KW Cushing's disease; dysmenorrhoea; antianginal; cytostatic; osteoporosis;
KW metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
KW tranquiliser; antiulcer; antiaddictive; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   27..1274
FT   /*tag= a
FT   /product= "Human SNORF72 receptor"
XX
PN WO200144297-A1.
XX
PD 21-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-US33787.
XX
PR 17-DEC-1999; 99US-0466435.
PR 25-APR-2000; 2000US-0558099.
PR 30-JUN-2000; 2000US-0609146.
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
XX
PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
XX
DR WPI; 2001-390240/41.
DR P-PSDB; AAE03629.
XX
PT A purified mammalian SNORF62 or SNORF72 receptor protein for
PT identification of compounds to treat e.g. inflammation, arthritis,
PT autoimmune diseases, transplant rejection, AIDS, cancer -
XX
PS Claim 42; Fig 3; 256pp; English.
XX
CC The invention relates to human G-protein coupled receptors, SNORF62 and
CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
CC are also known as NMU receptors. The agonist and antagonist of NMU
CC receptors are useful for treating an abnormality in a subject that is
CC alleviated by decreasing or increasing the activity of NMU receptor.
CC The NMU receptors serves as a valuable tool for designing drugs which are
CC useful for treating various pathophysiological conditions such as
CC inflammation, arthritis, autoimmune diseases, transplant rejection,
CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC septicaemia, AIDS, pain, psychotic and neurological disorders, including
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC eating/body weight disorders including obesity, bulimia, diabetes,
CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC disorders, ischaemia, stroke, cancers, sexual disorders, circadian
CC disorders, renal disorders, bone diseases including osteoporosis, benign
CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC dermatological disorders such as psoriasis, allergies, Parkinson's
CC disease, Alzheimer's disease, acute heart failure, angina disorders,
CC delirium and dyskinesias such as Huntington's disease. They can also be
CC used to regulate steroid hormone disorders, epinephrine release
CC disorders, electrolyte balance disorders, endocrine disorders, memory
CC disorders, somatosensory disorders, metabolic disorders, behavioural
CC disorders, drug addiction, migraine, Addison's disease, Cushing's
CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
CC The present cDNA sequence encodes human G-protein coupled receptor,
CC SNORF72.
XX
SQ Sequence 1298 BP; 278 A; 405 C; 282 G; 333 T; 0 other;

  Query Match   98.5%; Score 717.8; DB 22; Length 1298;
  Best Local Similarity   99.7%; Pred. No. 1.2e-145;
  Matches 719; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      36 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 95

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 155

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     156 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 215

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    216 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 275

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 335

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 395

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 455

      421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      456 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 515

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    516 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 575

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    576 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 635

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    636 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 695

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    696 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 755

Qy    721 A 721
          |
Db    756 A 756
```

RESULT 12
AAD23701
ID AAD23701 standard; cDNA; 1344 BP.
XX
AC AAD23701;
XX
DT 07-MAR-2002 (first entry)
XX
DE Human NMUR2 cDNA.
XX
KW Human; neuromedin U receptor; NMUR2; FM-4 protein; eating; weight gain;
KW feeding behaviour; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 55..1302
FT /*tag= a
FT /product= "Human NMUR2 protein"
XX
PN WO200181418-A1.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US13386.
XX
PR 27-APR-2000; 2000US-200718P.
XX
PA (MERI) MERCK & CO INC.
XX
PI Liu Q, Lynch KR, Howard AD, Mellin TN, Strack A;
PI Van Der Ploeg LHT, Wang R, Jiang Q, Williams D;
XX
DR WPI; 2002-034435/04.
DR P-PSDB; AAE14262.
XX
PT New polypeptide, useful for identifying compounds which modulate the
PT feeding activity of a mammal, comprises the novel isolated human and
PT rat neuromedin U receptor designated NMUR2 -
XX
PS Example 1; Fig 1; 47pp; English.
XX
CC The invention relates to human and rat neuromedin U receptor designated
CC NMUR2 polypeptides and polynucleotides. NMUR2 also referred as FM-4, is
CC free from associated proteins and is involved in the feeding behaviour
CC in mammals. Sequences of the invention are useful for identifying
CC compounds which modulate the feeding activity of a mammal. The compounds
CC identified are useful for modulating eating and weight gain. The present
CC sequence is human NMUR2 cDNA.
XX
SQ Sequence 1344 BP; 291 A; 419 C; 290 G; 344 T; 0 other;

```
  Query Match             98.5%;  Score 717.8;  DB 24;  Length 1344;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784
```

```
RESULT 13
AAZ33297
ID   AAZ33297 standard; cDNA; 1360 BP.
XX
AC   AAZ33297;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Human neurotensin-like receptor encoding cDNA.
XX
KW   Human; neurotensin-like receptor; NLR: G-protein coupled receptor;
KW   central nervous system; anasthesia; analgesia; ds.
XX
OS   Homo sapiens.
XX
PN   WO9955732-A1.
XX
PD   04-NOV-1999.
XX
PF   15-APR-1999;   99WO-SE00598.
XX
PR   24-APR-1998;   98SE-0001455.
XX
PA   (ASTR-) ASTRA PHARMA INC.
PA   (ASTR ) ASTRA AB.
XX
PI   Ahmad S,  Cao J,  O'Donnell D,  Walker P;
XX
DR   WPI; 2000-052803/04.
DR   P-PSDB; AAY52992.
XX
PT   Novel neurotensin-like receptor, useful for identifying agents for
PT   producing anasthesia or analgesia  -
XX
PS   Claim 11; Fig 1; 46pp; English.
XX
CC   The present sequence encodes a human G-protein coupled receptor
CC   neurotensin-like receptor (NLR). The NLR polynucleotide and protein
CC   can be used to isolate compounds that bind, (ant)agonise or alter
CC   the activity or expression of the NLR. The NLR is a G-protein coupled
CC   receptor which is expressed in the central nervous system and shares
CC   homology with human neurotensin receptor. The receptors can be used in
CC   assays to identify agents for producing anasthesia and analgesia.
XX
SQ   Sequence 1360 BP; 293 A; 423 C; 294 G; 350 T; 0 other;

  Query Match            98.5%;  Score 717.8;  DB 21;  Length 1360;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      74 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 133

Qy    61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 193

Qy   121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 253

Qy   181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 313

Qy   241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 373

Qy   301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 433

Qy   361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 493

Qy   421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 553

Qy   481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 613

Qy   541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   614 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 673

Qy   601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   674 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 733

Qy   661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   734 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 793

Qy   721 A 721
         |
Db   794 A 794


RESULT 14
AAF80326
ID   AAF80326 standard; DNA; 1594 BP.
XX
AC   AAF80326;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Splice variant of G-protein coupled receptor IGS4A long version cDNA.
XX
KW   Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW   nervous system disorder; psychiatric disorder; Parkinson's disease;
KW   episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW   cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW   motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW   gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW   inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW   gynecological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             55..945
FT                   /*tag= a
FT                   /product= "splice variant of G-protein coupled receptor
FT                            IGS4A long version"
XX
PN   WO200125269-A2.
XX
PD   12-APR-2001.
XX
PF   25-SEP-2000; 2000WO-EP09584.
XX
PR   24-SEP-1999;   99EP-0203140.
PR   24-SEP-1999;   99NL-1013140.
PR   28-JUL-2000; 2000EP-0202683.
PR   31-JUL-2000; 2000US-0222047.
XX
PA   (SOLV ) SOLVAY PHARM BV.
XX
PI   Deleersnijder W,  Berger C,  Loeken C,  Nys G,  Venema J;
```

```
XX
DR   WPI; 2001-273568/28.
DR   P-PSDB; AAB67806.
XX
PT   New G-protein coupled receptors and the polynucleotides encoding them,
PT   useful for preventing, ameliorating or correcting nervous system
PT   disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT   or cancers  -
XX
PS   Example 1b; Page 90-92; 102pp; English.
XX
CC   The present sequence encodes a splice variant of the long version of a
CC   human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC   polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
     polynucleotides are useful for preventing, ameliorating or correcting
     dysfunctions or diseases. These diseases include peripheral nervous
     system, psychiatric and central nervous system disorders
CC   (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC   disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC   (e.g. heart failure, angina pectoris, myocardial infarction or
CC   hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC   disorders (e.g. inflammatory bowel disease or motility disorders),
CC   osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC   protozoan or viral), pain, cancers, immune disorders, allergies,
CC   sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC   effective with regard to disorders of the nervous system, including the
CC   central and peripheral nervous systems, disorders of the gastrointestinal
CC   system, cardiovascular system, skeletal muscle, thyroid, lung or
CC   genitourinary system, or immunological disease. The IGS4 polynucleotides
CC   are useful as diagnostic reagents for detecting under-expression,
CC   overexpression or altered expression of IGS4.
XX
SQ   Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other;

  Query Match             98.5%;  Score 717.8;  DB 22;  Length 1594;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy     721 A 721
          |
Db     784 A 784


RESULT 15
AAF80327
ID   AAF80327 standard; DNA; 1594 BP.
XX
AC   AAF80327;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Splice variant of G-protein coupled receptor IGS4A short version cDNA.
XX
KW   Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW   nervous system disorder; psychiatric disorder; Parkinson's disease;
KW   episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW   cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW   motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW   gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW   inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW   gynecological disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             64..945
FT                   /*tag=  a
FT                   /product= "splice variant of G-protein coupled receptor
FT                            IGS4A short version"
XX
PN   WO200125269-A2.
XX
PD   12-APR-2001.
XX
PF   25-SEP-2000; 2000WO-EP09584.
XX
PR   24-SEP-1999;   99EP-0203140.
PR   24-SEP-1999;   99NL-1013140.
PR   28-JUL-2000; 2000EP-0202683.
PR   31-JUL-2000; 2000US-0222047.
XX
PA   (SOLV ) SOLVAY PHARM BV.
XX
PI   Deleersnijder W,  Berger C,  Loeken C,  Nys G,  Venema J;
XX
DR   WPI; 2001-273568/28.
DR   P-PSDB; AAB67807.
XX
PT   New G-protein coupled receptors and the polynucleotides encoding them,
PT   useful for preventing, ameliorating or correcting nervous system
PT   disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT   or cancers  -
XX
PS   Example 1b; Page 94-95; 102pp; English.
XX
CC   The present sequence encodes a splice variant of the short version of a
CC   human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC   polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
```

```
CC   polynucleotides are useful for preventing, ameliorating or correcting
CC   dysfunctions or diseases. These diseases include peripheral nervous
CC   system, psychiatric and central nervous system disorders
CC   (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC   disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC   (e.g. heart failure, angina pectoris, myocardial infarction or
CC   hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC   disorders (e.g. inflammatory bowel disease or motility disorders),
CC   osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC   protozoan or viral), pain, cancers, immune disorders, allergies,
CC   sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC   effective with regard to disorders of the nervous system, including the
CC   central and peripheral nervous systems, disorders of the gastrointestinal
CC   system, cardiovascular system, skeletal muscle, thyroid, lung or
CC   genitourinary system, or immunological disease. The IGS4 polynucleotides
CC   are useful as diagnostic reagents for detecting under-expression,
CC   overexpression or altered expression of IGS4.
XX
SQ   Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other;

  Query Match             98.5%;  Score 717.8;  DB 22;  Length 1594;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-145;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

     1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784


Search completed: January 17, 2003, 02:34:05
Job time : 272 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:34:16 ; Search time 2748 Seconds
(without alignments)
7720.504 Million cell updates/sec

Title: US-09-684-725-1
Perfect score: 729
Sequence: 1 atggaaaaacttcagaatgc.....cactcagagtgagtatctag 729

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		729	100.0	729	6	AX139107	AX139107 Sequence
2	c	729	100.0	801	6	AX147772	AX147772 Sequence
3	c	729	100.0	214267	9	AC008571	AC008571 Homo sapi
4		717.8	98.5	1239	9	AF242874	AF242874 Homo sapi
5		717.8	98.5	1248	9	AB041228	AB041228 Homo sapi
6		717.8	98.5	1298	9	AF272363	AF272363 Homo sapi
7		717.8	98.5	1594	6	AX109242	AX109242 Sequence
8		717.8	98.5	1594	6	AX109244	AX109244 Sequence
9		717.8	98.5	1658	6	AX109234	AX109234 Sequence
10		717.8	98.5	1658	6	AX109236	AX109236 Sequence
11		717.8	98.5	1658	6	AX109238	AX109238 Sequence
12		717.8	98.5	1658	6	AX109240	AX109240 Sequence
13		716.2	98.2	1248	9	AF292402	AF292402 Homo sapi
14		716.2	98.2	2069	9	BC016938	BC016938 Homo sapi
15	c	476.8	65.4	165392	2	AC122979	AC122979 Rattus no
16	c	476.8	65.4	178043	2	AC110442	AC110442 Rattus no
17	c	476.2	65.3	213942	2	AL662792	AL662792 Mus muscu
18		471.4	64.7	1188	10	AF242875	AF242875 Rattus no
19		471.4	64.7	1188	10	AY057384	AY057384 Mus muscu
20		471.4	64.7	1314	10	AB041229	AB041229 Rattus no
21		282.4	38.7	3354	9	BC036543	BC036543 Homo sapi
22		280.8	38.5	828	9	HSOGPCR1	AF044600 Homo sapi
23		280.8	38.5	1209	6	BD012783	BD012783 Screening
24		280.8	38.5	1209	23	BD008218	Bd008218 Screening
25		280.8	38.5	1318	9	AF272362	AF272362 Homo sapi
26	c	280.8	38.5	168880	9	AC017104	AC017104 Homo sapi
27		277.6	38.1	2374	6	AX342663	AX342663 Sequence
28		255	35.0	1209	10	AF242873	AF242873 Rattus no
29		255	35.0	1239	10	AB038649	AB038649 Rattus no
30		255	35.0	171749	2	AC112440	AC112440 Rattus no
31		243	33.3	1218	10	AF044602	AF044602 Mus muscu
32	c	243	33.3	203177	2	AC102609	AC102609 Mus muscu
33		161.6	22.2	167102	2	AC118178	AC118178 Rattus no
34	c	159.6	21.9	75950	2	AC073449	AC073449 Homo sapi
35		134.8	18.5	1161	6	AX154591	AX154591 Sequence
36		134.8	18.5	1239	6	AX154589	AX154589 Sequence
37		134.8	18.5	2040	9	AF034632	AF034632 Homo sapi
38		134.8	18.5	163284	9	AL137000	AL137000 Human DNA
39		134.8	18.5	341560	2	AL596304	AL596304 Homo sapi
40		133.2	18.3	1676	5	AF082210	AF082210 Sphoeroid
41		131.2	18.0	692	9	HSA339407	AJ339407 Homo sapi
42		127	17.4	1254	6	E11480	E11480 cDNA encodi
43		127	17.4	1257	6	AX280911	AX280911 Sequence
44		127	17.4	4131	9	HSNEURA	X70070 H.sapiens m
45		125.8	17.3	615	9	HSA339459	AJ339459 Homo sapi

ALIGNMENTS

RESULT 1
AX139107
LOCUS AX139107 729 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090990.
ACCESSION AX139107
VERSION AX139107.1 GI:14274786
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS Harland,L.P.
TITLE Human g-protein-coupled receptor
JOURNAL Patent: EP 1090990-A 1 11-APR-2001;

```
           Pfizer Limited (GB) ; PFIZER INC. (US)
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      135 a    245 c    167 g    182 t
ORIGIN

  Query Match             100.0%;  Score 729;  DB 6;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-127;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180

       181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720

Qy     721 AGTATCTAG 729
           |||||||||
Db     721 AGTATCTAG 729


RESULT 2
AX147772/c
LOCUS       AX147772                 801 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 17 from Patent WO0136473.
ACCESSION   AX147772
VERSION     AX147772.1  GI:14346807
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
            Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
            Sejlitz,T. and Huff,R.M.
  TITLE     Novel g protein-coupled receptors
  JOURNAL   Patent: WO 0136473-A 17 25-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      200 a    187 c    261 g    153 t
ORIGIN

  Query Match             100.0%;  Score 729;  DB 6;  Length 801;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-127;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 736

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 676

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     675 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 616

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 556

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 496

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 436

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 376

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 316

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 256

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 196

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 136

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 76

Qy     721 AGTATCTAG 729
           |||||||||
Db      75 AGTATCTAG 67
```

```
RESULT 3
AC008571/c
LOCUS       AC008571              214267 bp    DNA     linear   PRI 26-SEP-2001
DEFINITION  Homo sapiens chromosome 5 clone CTC-550M4, complete sequence.
ACCESSION   AC008571
VERSION     AC008571.6  GI:15778683
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214267)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214267)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 214267)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 26, 2001 this sequence version replaced gi:15290296.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of Sequence;
            Estimated Total Number of Errors is 0.3.
FEATURES             Location/Qualifiers
     source          1. .214267
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-550M4"
BASE COUNT    62950 a  40311 c  43061 g  67945 t
ORIGIN

Query Match           100.0%;  Score 729;  DB 9;  Length 214267;
Best Local Similarity  100.0%;  Pred. No. 1.4e-127;
Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104727 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 104668

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104667 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 104608

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104607 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 104548

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104547 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 104488

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104487 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 104428

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104427 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 104368

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104367 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 104308

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104307 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 104248

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104247 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 104188

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104187 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 104128

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104127 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 104068

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104067 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 104008

Qy    721 AGTATCTAG 729
          |||||||||
Db 104007 AGTATCTAG 103999


RESULT 4
AF242874
LOCUS       AF242874               1239 bp    mRNA    linear   PRI 13-JUL-2000
DEFINITION  Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.
ACCESSION   AF242874
VERSION     AF242874.1  GI:9082155
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1239)
  AUTHORS   Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M.,
            Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B.,
            Stair,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K.,
            Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P.,
            Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
  TITLE     Identification of receptors for neuromedin U and its role in
            feeding
  JOURNAL   Nature 406 (6791), 70-74 (2000)
  MEDLINE   20351041
   PUBMED   10894543
REFERENCE   2  (bases 1 to 1239)
  AUTHORS   Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
            Point, PA 19486, USA
FEATURES             Location/Qualifiers
     source          1. .1239
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     gene            1. .1239
                     /gene="NMU2R"
     CDS             1. .1239
                     /gene="NMU2R"
                     /note="FM-4"
                     /codon_start=1
                     /product="neuromedin U receptor 2"
                     /protein_id="AAF82755.1"
                     /db_xref="GI:9082156"
                     /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
                     VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
                     MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
                     LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
                     FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI
                     CWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRFQAA
```

FQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSHLPT
ALSSEQMSRTNYQSFHFNKT"
BASE COUNT 267 a 392 c 265 g 315 t
ORIGIN

Query Match 98.5%; Score 717.8; DB 9; Length 1239;
Best Local Similarity 99.7%; Pred. No. 2.2e-125;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120

Qy     121 TTCTTCCTCCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 720

Qy     721 A 721
          |
Db     721 A 721
```

RESULT 5
AB041228
LOCUS AB041228 1248 bp mRNA linear PRI 21-SEP-2000
DEFINITION Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.
ACCESSION AB041228
VERSION AB041228.1 GI:10257380
KEYWORDS G protein-coupled receptor TGR-1.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shintani,Y., Fukusumi,S., Habata,Y., Hinuma,S., Onda,H., Nishimura,O. and Fujino,M.
TITLE Identification and functional characterization of a novel subtype of neuromedin U receptor
JOURNAL J. Biol. Chem. 275 (38), 29528-29532 (2000)
MEDLINE 20449029
REFERENCE 2 (bases 1 to 1248)
AUTHORS Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.
TITLE Direct Submission
JOURNAL Submitted (03-APR-2000) Yasushi Shintani, Takeda Chemical Industries, LTD, Discovery Research Laboratories 1; 10 Wadai, Tsukuba, Ibaraki 300-4293, Japan
(E-mail:Shintani_Yasushi@takeda.co.jp, Tel:81-298-64-5011, Fax:81-298-64-5000)
FEATURES Location/Qualifiers
source 1. .1248
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .1248
/gene="TGR-1"
CDS 10. .1248
/gene="TGR-1"
/codon_start=1
/product="G protein-coupled receptor TGR-1"
/protein_id="BAB13721.1"
/db_xref="GI:10257381"
/translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI
CWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRFQAA
FQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSHLPT
ALSSEQMSRTNYQSFHFNKT"
BASE COUNT 269 a 393 c 269 g 317 t
ORIGIN

Query Match 98.5%; Score 717.8; DB 9; Length 1248;
Best Local Similarity 99.7%; Pred. No. 2.2e-125;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

          661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy    721 A 721
          |
Db    730 A 730
```

RESULT 6
AF272363
LOCUS AF272363 1298 bp mRNA linear PRI 23-OCT-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.
ACCESSION AF272363
VERSION AF272363.1 GI:10946202
KEYWORDS .
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1298)
AUTHORS Raddatz,R., Wilson,A.E., Artymyshyn,R., Bonini,J.A., Borowsky,B., Boteju,L.W., Zhou,S., Kouranova,E.V., Nagorny,R., Guevarra,M.S., Dai,M., Lerman,G.S., Vaysse,P.J., Branchek,T.A., Gerald,C., Forray,C. and Adham,N.
TITLE Identification and characterization of two neuromedin U receptors differentially expressed in peripheral tissues and the central nervous system
JOURNAL J. Biol. Chem. 275 (42), 32452-32459 (2000)
MEDLINE 20490668
PUBMED 10899166
REFERENCE 2 (bases 1 to 1298)
AUTHORS Bonini,J.A., Raddatz,R., Wilson,A. and Borowsky,B.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652, USA
FEATURES Location/Qualifiers
source 1. .1298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q31.1-q31.3"
gene 1. .1298
/gene="NMUR2"
CDS 27. .1274
/gene="NMUR2"
/note="G protein-coupled receptor"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AAG24794.1"
/db_xref="GI:10946203"
/translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQ
VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRF
QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
LPTALSSEQMSRTNYQSFHFNKT"
BASE COUNT 278 a 405 c 282 g 333 t
ORIGIN

Query Match 98.5%; Score 717.8; DB 9; Length 1298;
Best Local Similarity 99.7%; Pred. No. 2.2e-125;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 95

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 155

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 215

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    216 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 275

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 335

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 395

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 455

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 515

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    516 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 575

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    576 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 635

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    636 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 695

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    696 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 755

Qy    721 A 721
          |
Db    756 A 756
```

RESULT 7
AX109242
LOCUS AX109242 1594 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 9 from Patent WO0125269.
ACCESSION AX109242
VERSION AX109242.1 GI:13924115
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 9 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             55. .945
                     /note="IGS4A truncated DNA long version"
                     /codon_start=1
                     /protein_id="CAC37775.1"
                     /db_xref="GI:13924116"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQ
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLSLWRSG
                     VNPWLLCSTSSMWCQVSSST"
BASE COUNT      370 a    473 c    332 g    419 t
ORIGIN

  Query Match             98.5%;  Score 717.8;  DB 6;  Length 1594;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |

Db     724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy     721 A 721
           |
Db     784 A 784


RESULT 8
AX109244
LOCUS       AX109244               1594 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 11 from Patent WO0125269.
ACCESSION   AX109244
VERSION     AX109244.1  GI:13924117
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 11 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             64. .945
                     /note="IGS4A truncated DNA short version"
                     /codon_start=1
                     /protein_id="CAC37776.1"
                     /db_xref="GI:13924118"
                     /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
                     VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
                     MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
                     LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
                     FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLSLWRSGVNP
                     WLLCSTSSMWCQVSSST"
BASE COUNT      370 a    473 c    332 g    419 t
ORIGIN

  Query Match             98.5%;  Score 717.8;  DB 6;  Length 1594;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483
```

```
Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784
```

```
RESULT 9
AX109234
LOCUS       AX109234                1658 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 1 from Patent WO0125269.
ACCESSION   AX109234
VERSION     AX109234.1  GI:13924107
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1658)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 1 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             55. .1302
                     /note="IGS4A long version"
                     /codon_start=1
                     /protein_id="CAC37771.1"
                     /db_xref="GI:13924108"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQ
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
                     FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRF
                     QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
                     LPTALSSEQMSRTNYQSFHFNKT"
BASE COUNT      377 a    490 c    346 g    445 t
ORIGIN

  Query Match             98.5%;  Score 717.8;  DB 6;  Length 1658;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183
```

```
Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784
```

```
RESULT 10
AX109236
LOCUS       AX109236                1658 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 3 from Patent WO0125269.
ACCESSION   AX109236
VERSION     AX109236.1  GI:13924109
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1658)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 3 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             64. .1302
                     /note="IGS4A short version"
                     /codon_start=1
                     /protein_id="CAC37772.1"
                     /db_xref="GI:13924110"
                     /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
                     VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
                     MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
```

```
                    LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
                    FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI
                    CWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRFQAA
                    FQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSHLPT
                    ALSSEQMSRTNYQSFHFNKT"
BASE COUNT      377 a    490 c    346 g    445 t
ORIGIN

  Query Match             98.5%;  Score 717.8;  DB 6;  Length 1658;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

       181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy     721 A 721
          |
Db     784 A 784


RESULT 11
AX109238
LOCUS       AX109238               1658 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 5 from Patent WO0125269.
ACCESSION   AX109238
VERSION     AX109238.1  GI:13924111
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1658)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 5 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             55. .1302
                     /note="IGS4B long version"
                     /codon_start=1
                     /protein_id="CAC37773.1"
                     /db_xref="GI:13924112"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQ
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
                     FAICWAPFHIDRLFFSFVEEWTESLAAVFNLVHVVSGVLFYLSSAVNPIIYNLLSRRF
                     QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFLCQSSVHNSH
                     LPTALSSEQMSRTNYQSFHFNKT"
BASE COUNT      378 a    489 c    346 g    445 t
ORIGIN

  Query Match             98.5%;  Score 717.8;  DB 6;  Length 1658;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
```

```
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784


RESULT 12
AX109240
LOCUS       AX109240               1658 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 7 from Patent WO0125269.
ACCESSION   AX109240
VERSION     AX109240.1  GI:13924113
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1658)
  AUTHORS   Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
  TITLE     Human g-protein coupled receptor
  JOURNAL   Patent: WO 0125269-A 7 12-APR-2001;
            Solvay Pharmaceuticals B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .1658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             64. .1302
                     /note="IGS4B short version"
                     /codon_start=1
                     /protein_id="CAC37774.1"
                     /db_xref="GI:13924114"
                     /translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
                     VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
                     MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
                     LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
                     FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI
                     CWAPFHIDRLFFSFVEEWTESLAAVFNLVHVVSGVLFYLSSAVNPIIYNLLSRRFQAA
                     FQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFLCQSSVHNSHLPT
                     ALSSEQMSRTNYQSFHFNKT"
BASE COUNT      378 a    489 c    346 g    445 t
ORIGIN

  Query Match           98.5%;  Score 717.8;  DB 6;  Length 1658;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-125;
  Matches  719;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 123

Qy     61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 183

Qy    121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 243

Qy    181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 303

Qy    241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 363

Qy    301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 423

Qy    361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 483

Qy    421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 543

Qy    481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 603

Qy    541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 663

Qy    601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 723

Qy    661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    724 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 783

Qy    721 A 721
          |
Db    784 A 784


RESULT 13
AF292402
LOCUS       AF292402               1248 bp    mRNA    linear   PRI 11-DEC-2000
DEFINITION  Homo sapiens neuromedin U receptor-type 2 mRNA, complete cds.
ACCESSION   AF292402
VERSION     AF292402.1  GI:9944989
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1248)
  AUTHORS   Shan,L., Qiao,X., Crona,J.H., Behan,J., Wang,S., Laz,T., Bayne,M.,
            Gustafson,E.L., Monsma,F.J. Jr. and Hedrick,J.A.
  TITLE     Identification of a novel neuromedin U receptor subtype expressed
            in the central nervous system
  JOURNAL   J. Biol. Chem. 275 (50), 39482-39486 (2000)
  MEDLINE   20564321
   PUBMED   11010960
REFERENCE   2  (bases 1 to 1248)
  AUTHORS   Pang,L., Wang,S., Laz,T. and Hedrick,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2000) Human Genome Research, Schering-Plough
            Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033,
            USA
FEATURES             Location/Qualifiers
     source          1. .1248
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
     CDS             1. .1248
                     /note="G protein coupled receptor; NmU-R2"
                     /codon_start=1
                     /product="neuromedin U receptor-type 2"
                     /protein_id="AAG03064.1"
                     /db_xref="GI:9944990"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQ
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
```

```
                     FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRF
                     QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
                     LPTALSSEQMSRTNYQSFHFNKT"
BASE COUNT      268 a    393 c    270 g    317 t
ORIGIN

  Query Match             98.2%;  Score 716.2;  DB 9;  Length 1248;
  Best Local Similarity   99.6%;  Pred. No. 4.5e-125;
  Matches  718;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      10 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCGTTCCAG 69

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 129

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 189

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 249

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 309

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     310 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 369

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 429

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 489

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 549

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     550 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 609

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 669

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     670 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 729

Qy     721 A 721
           |
Db     730 A 730


RESULT 14
BC016938
LOCUS       BC016938               2069 bp    mRNA    linear   PRI 09-NOV-2001
DEFINITION  Homo sapiens, neuromedin U receptor 2, clone MGC:21396
            IMAGE:3852151, mRNA, complete cds.
ACCESSION   BC016938
VERSION     BC016938.1  GI:16877376
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2069)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 20 Row: n Column: 3
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .2069
                     /organism="Homo sapiens"
                     /db_xref="LocusID:56923"
                     /db_xref="taxon:9606"
                     /clone="MGC:21396 IMAGE:3852151"
                     /tissue_type="Colon, adenocarcinoma"
                     /clone_lib="NIH_MGC_65"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             162. .1409
                     /codon_start=1
                     /product="neuromedin U receptor 2"
                     /protein_id="AAH16938.1"
                     /db_xref="GI:16877377"
                     /translation="MSGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHF
                     FLPVSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLE
                     VYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTR
                     RRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATFTVIKPMWIYNFIIQ
                     VTSFLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLV
                     FAICWAPFHIDRLFFSFVEEWSESLAAVFNLVHVVSGVFFYLSSAVNPIIYNLLSRRF
                     QAAFQNVISSFHKQWHSQHDPQLPPAQRNIFLTECHFVELTEDIGPQFPCQSSMHNSH
                     LPTALSSEQMSRTNYQSFHFNKT"
BASE COUNT      481 a    579 c    411 g    598 t
ORIGIN

  Query Match             98.2%;  Score 716.2;  DB 9;  Length 2069;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-125;
  Matches  718;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 230

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 290

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 350

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 410
```

```
Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 470

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 530

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     531 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 590

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       591 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 650

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     651 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 710

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     711 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 770

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     771 TTTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 830

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     831 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGACTA 890

Qy     721 A 721
           |
Db     891 A 891
```

```
RESULT 15
AC122979/c
LOCUS       AC122979              165392 bp    DNA     linear   HTG 24-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-353A10, *** SEQUENCING IN PROGRESS
            ***, 59 unordered pieces.
ACCESSION   AC122979
VERSION     AC122979.2  GI:21909359
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 165392)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 165392)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 165392)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 19, 2002 this sequence version replaced gi:21217312.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GYOC
                Center clone name: CH230-353A10
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 121972 bases at least Q40
                Consensus quality: 127005 bases at least Q30
                Consensus quality: 130275 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 59 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1383: contig of 1383 bp in length
            *    1384     1483: gap of unknown length
            *    1484     2655: contig of 1172 bp in length
            *    2656     2755: gap of unknown length
            *    2756     4076: contig of 1321 bp in length
            *    4077     4176: gap of unknown length
            *    4177     5570: contig of 1394 bp in length
            *    5571     5670: gap of unknown length
            *    5671     7249: contig of 1579 bp in length
            *    7250     7349: gap of unknown length
            *    7350     8683: contig of 1334 bp in length
            *    8684     8783: gap of unknown length
            *    8784     9787: contig of 1004 bp in length
```

```
*     9788     9887: gap of unknown length
*     9888    11239: contig of 1352 bp in length
*    11240    11339: gap of unknown length
*    11340    12727: contig of 1388 bp in length
*    12728    12827: gap of unknown length
*    12828    14694: contig of 1867 bp in length
*    14695    14794: gap of unknown length
*    14795    16261: contig of 1467 bp in length
*    16262    16361: gap of unknown length
*    16362    17717: contig of 1356 bp in length
*    17718    17817: gap of unknown length
*    17818    19291: contig of 1474 bp in length
*    19292    19391: gap of unknown length
*    19392    20650: contig of 1259 bp in length
*    20651    20750: gap of unknown length
*    20751    23327: contig of 2577 bp in length
*    23328    23427: gap of unknown length
*    23428    24665: contig of 1238 bp in length
*    24666    24765: gap of unknown length
*    24766    26126: contig of 1361 bp in length
*    26127    26226: gap of unknown length
*    26227    27240: contig of 1014 bp in length
*    27241    27340: gap of unknown length
*    27341    29397: contig of 2057 bp in length
*    29398    29497: gap of unknown length
*    29498    30863: contig of 1366 bp in length
*    30864    30963: gap of unknown length
*    30964    32601: contig of 1638 bp in length
*    32602    32701: gap of unknown length
*    32702    35147: contig of 2446 bp in length
*    35148    35247: gap of unknown length
*    35248    36836: contig of 1589 bp in length
*    36837    36936: gap of unknown length
*    36937    38542: contig of 1606 bp in length
*    38543    38642: gap of unknown length
*    38643    40461: contig of 1819 bp in length
*    40462    40561: gap of unknown length
*    40562    41824: contig of 1263 bp in length
*    41825    41924: gap of unknown length
*    41925    44599: contig of 2675 bp in length
*    44600    44699: gap of unknown length
*    44700    46495: contig of 1796 bp in length
*    46496    46595: gap of unknown length
*    46596    48924: contig of 2329 bp in length
*    48925    49024: gap of unknown length
*    49025    51293: contig of 2269 bp in length
*    51294    51393: gap of unknown length
*    51394    53670: contig of 2277 bp in length
*    53671    53770: gap of unknown length
*    53771    56177: contig of 2407 bp in length
*    56178    56277: gap of unknown length
*    56278    57982: contig of 1705 bp in length
*    57983    58082: gap of unknown length
*    58083    60731: contig of 2649 bp in length
*    60732    60831: gap of unknown length
*    60832    63615: contig of 2784 bp in length
*    63616    63715: gap of unknown length
*    63716    66396: contig of 2681 bp in length
*    66397    66496: gap of unknown length
*    66497    69328: contig of 2832 bp in length
*    69329    69428: gap of unknown length
*    69429    72236: contig of 2808 bp in length
*    72237    72336: gap of unknown length
*    72337    74916: contig of 2580 bp in length
*    74917    75016: gap of unknown length
*    75017    77675: contig of 2659 bp in length
*    77676    77775: gap of unknown length
*    77776    81127: contig of 3352 bp in length
*    81128    81227: gap of unknown length
*    81228    85820: contig of 4593 bp in length
*    85821    85920: gap of unknown length
*    85921    88975: contig of 3055 bp in length
*    88976    89075: gap of unknown length
*    89076    92862: contig of 3787 bp in length
*    92863    92962: gap of unknown length
*    92963    95973: contig of 3011 bp in length
*    95974    96073: gap of unknown length
*    96074   100741: contig of 4668 bp in length
*   100742   100841: gap of unknown length
*   100842   103953: contig of 3112 bp in length
*   103954   104053: gap of unknown length
*   104054   106584: contig of 2531 bp in length
*   106585   106684: gap of unknown length
*   106685   110379: contig of 3695 bp in length
*   110380   110479: gap of unknown length
*   110480   113599: contig of 3120 bp in length
*   113600   113699: gap of unknown length
*   113700   117699: contig of 4000 bp in length
*   117700   117799: gap of unknown length
*   117800   122249: contig of 4450 bp in length
*   122250   122349: gap of unknown length
*   122350   126688: contig of 4339 bp in length
*   126689   126788: gap of unknown length
*   126789   133046: contig of 6258 bp in length

 Query Match             65.4%;  Score 476.8;  DB 2;  Length 165392;
 Best Local Similarity   79.8%;  Pred. No. 4e-80;
 Matches  581;  Conservative    0;  Mismatches  132;  Indels   15;  Gaps    1;

Qy       1 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 60
           |||| ||||||| | ||||||||||||||| ||               |||||  ||  |
Db 129275 ATGGGAAAACTTGAAAATGCTTCCTGGATCCAC--------------GATCCTCTCATG 129231

Qy      61 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 120
           ||  || |||||||||| ||||||||  |||||  ||| || |||||   ||||||  ||
Db 129230 AAGTACTTGAACAGCACAGAGGAGTACTTGGCCCACCTGTGTGGACCCAAGCGCAGTGAC 129171

Qy     121 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 180
           | | ||| || ||||||||||  |||| ||  || ||  ||||||||||| || ||||||
Db 129170 CTATCCCTTCCGGTGTCTGTGGCCTATGCGCTGATCTTCCTGGTGGGGGTAATGGGCAAT 129111

Qy     181 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 240
           | |||||||||| |||||||| | |  || ||| || ||||||| ||||||||||||||
Db 129110 CTTCTGGTGTGCATGGTGATTGTCCGACATCAGACTTTGAAGACACCCACCAACTACTAT 129051

Qy     241 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 300
           ||||||||| |||| ||||| || || ||||||||||||| | || ||||| |||||  ||
Db 129050 CTCTTCAGCTTGGCAGTCTCAGATCTGCTGGTCCTGCTCTTGGGGATGCCTCTGGAAATC 128991

Qy     301 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 360
           || |||||||||| ||| |||||||||| |||||||||| ||||| ||||||||||||||
Db 128990 TACGAGATGTGGCACAATTACCCTTTCCTGTTCGGGCCTGTGGGATGCTACTTCAAGACA 128931

Qy     361 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 420
           |||||||| ||||| |||||||| |||||||| ||||| ||||||| || ||||| |||
Db 128930 GCCCTCTTCGAGACTGTGTGCTTTGCCTCCATTCTCAGTGTCACCACGGTTAGCGTAGAG 128871

Qy     421 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 480
           ||||| |||||||||  | ||||| ||||| ||||| ||| ||||||| ||||| ||||||
Db 128870 CGCTATGTGGCCATTGTCCACCCTTTCCGAGCCAAGCTGGAGAGCACGCGGCGACGGGCC 128811

Qy     481 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 540
           |||||||||||| || | |||||| ||||||| ||| |||| ||  ||||||| ||||||
Db 128810 CTCAGGATCCTCAGCCTAGTCTGGAGCTTCTCTGTGGTCTTTTCTTTGCCCAATACCAGC 128751

Qy     541 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 600
           |||||||||||||||||||  |||| ||||| ||||||   || || || || ||||||
Db 128750 ATCCATGGCATCAAGTTCCAGCACTTTCCCAACGGGTCCTCCGTACCTGGCTCAGCCACC 128691

Qy     601 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 660
           || || |||| ||| |||||||||| | || || || |||||||| |  |||  ||||||
Db 128690 TGCACAGTCACCAAACCCATGTGGGTGTATAACTTGATCATCCAAGCTACCAGCTTCCTC 128631

Qy     661 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 720
           |||||| ||||||| |||||  ||||||| |||||||||||||||||||  ||||| |||
Db 128630 TTCTACATCCTCCCAATGACCCTCATCAGCGTCCTCTACTACCTCATGGGGCTCAGGGTG 128571
```

```
Qy     721 AGTATCTA 728
           |||||| |
Db 128570 AGTATCGA 128563
```

Search completed: January 17, 2003, 04:30:06
Job time : 3043 secs

THIS PAGE BLANK (USPTO)